Bacterial sialic acid degradation: membrane transport and enzymology

---

A thesis submitted in partial fulfilment of the

requirements for the degree of

Doctor of Philosophy in Biochemistry

in the School of Biological Sciences

by Rachel Aimee North

University of Canterbury

May 2016

Abstract

The overarching aim of this thesis is to understand how sialic acids are transported into the bacterial cell and then degraded by bacterial pathogens. In heavily sialylated environments, bacterial pathogens utilise host-derived sialic acid as a nutrient source, and this pathway constitutes a novel and unexploited target for antibiotic drug design.

Three sialic acid transporters, from two distinct gene families were investigated to probe how sialic acid is transported across the cytoplasmic membrane. The *Yersinia pestis* sugar proton symporter (NanT) exists as a single species in solution, while the *Staphylococcus aureus* sodium solute symporter (SSS) and *Proteus mirabilis* SSS appear to be in a self-association. It is likely that the SSS sialic acid transporters exist in a monomer-dimer equilibrium, although higher oligomeric structures cannot be ruled out. The structure of the *P. mirabilis* SSS sialic acid transporter was solved, which is the first known structure of a sialic acid transporter to be presented. It was solved in complex with sialic acid and two sodium ions, providing insight into how this transporter mediates the movement of sialic acid across the membrane. In addition, the structure presents a novel conformation among sodium symporters and the structural basis for the movement of sialic acid across the membrane is elucidated.

Following the import of sialic acid into the bacterial cell, *N*-acetylneuraminate lyase is the first enzyme involved in its degradation. The structure, function and inhibition of methicillin-resistant *S. aureus* (MRSA) *N*-acetylneuraminate lyase were investigated. Solution and structural studies confirmed that this enzyme is tetrameric. Kinetic analysis was employed to test an inhibitor of *N*-acetylneuraminate lyase enzymes. This molecule demonstrated strong species-specific inhibition against MRSA *N*-acetylneuraminate lyase. The structure of MRSA *N*-acetylneuraminate lyase in complex with this inhibitor demonstrated that altered binding modes between *N*-acetylneuraminate lyase enzymes may account for variable inhibition between species. In addition, evidence for a protein-protein interaction between the *S. aureus* SSS sialic acid transporter and MRSA *N*-acetylneuraminate lyase is presented. Thus, these proteins may interact at the cytoplasmic membrane, allowing the degradation of sialic acid to be initiated upon entering the cell.

The structure, function and catalytic mechanism of the third enzyme involved in sialic acid degradation, *N*-acetylmannosamine-6-phosphate 2-epimerase, were explored from MRSA. This enzyme was demonstrated to adopt a dimeric architecture in solution, consistent with the crystal structure that was solved and presented. A multi-enzyme coupled assay was developed to assess *N*-acetylmannosamine-6-phosphate 2-epimerase activity in real time. Residues were probed that may be important for catalysis and tested by mutagenesis and kinetic analysis. A novel mechanism of carbohydrate epimerisation is proposed for this enzyme, whereby catalysis occurs *via* a proton displacement mechanism mediated by the substrate.

Overall, this work provides new data that enriches our understanding of the import and degradation of sialic acid in clinically important human bacterial pathogens.

Acknowledgements

My biggest thank you is owed to my supervisor, Renwick Dobson, whose enthusiasm and passion for science is inspiring. Thanks for all of the incredible opportunities and for making the last four years enjoyable. But most of all, thank you for your unwavering support, guidance and encouragement. Thanks also to the other members of my PhD committee; Sarah Kessans, Andrew Muscroft-Taylor and Juliet Gerrard for your interest, advice and input into my project.

Thanks to Richard Neutze for inviting me to work on part of my project in your lab, which unknowingly was the exact same project as the group next door to you. Thus, an incredibly big thank you goes to Rosmarie Friemann, for allowing me to work so closely with your group. Thanks to all of your group members for sharing the project and for teaching me.

A special thank you goes to Kat, for your friendship and support during these challenging four years. I don't know how I would have got through it without you. Thank you to all the people that have been a part of the Dobson, Gerrard and Pearce labs over the last four years. In particular, thanks to Jen, Chris, Jeremy, Amy and Arvind. I feel privileged to have worked with such an amazing group of people and to form such great friendships with you all. Thank you to Grant for your graphical assistance and thanks to Jackie for making sure the lab was always running smoothly.

Thank you to the various organisations that provided a significant amount of financial support over the years; to the University of Canterbury for a Doctoral Scholarship, the European Molecular Biology Organisation for a Short-Term Fellowship and to the New Zealand Synchrotron Group for a scholarship to attend the 2014 Cheiron School at SPring-8 in Japan. To the New Zealand Society for Biochemistry and Molecular Biology, the New Zealand Federation of Graduate Women, the Royal Society of New Zealand, the Protein Society, the Maurice and Phyllis Paykel Trust and the New Zealand Institute of Chemistry for funding various trips to Sweden and conferences all over the world. Thanks to the Biomolecular Interaction Centre for the provision of equipment and various sources of expert advice.

Thank you to my incredibly supportive friends outside of the lab, whom always showed an interest in my research and praised all of my achievements. Thank you to Mum, Dad and Laura for your love, support, patience and encouragement; I truly could not have done it without you.

Table of contents

Abbreviations

ADP	adenosine diphosphate
ATP	adenosine triphosphate
ABC	adenosine triphosphate binding cassette transporter
BetP	sodium glycine betaine symporter
BLAST	basic local alignment search tool
C3	Collaborative Crystallisation Centre
CD	circular dichroism
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
cm	centimetre
c(*M*)	continuous mass distribution
c(*s*)	continuous sedimentation coefficient distribution
CMP	cytidine monophosphate
Da	dalton
DDM	n-dodecyl-β-D-maltoside
DHDPS	dihydrodipicolinate synthase
DM	n-decyl-β-D-maltoside
D_{max}	maximum particle diameter
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
EDTA	ethylenediaminetetraacetic acid
EMBO	European Molecular Biology Organisation
F	fluorescence
f/f_0	frictional ratio
F_N	normalised fluorescence
FSEC	fluorescence detection size exclusion chromatography
g	gram
GFP	green fluorescent protein
His-tag	Histidine tag
^{1}H NMR	proton nuclear magnetic resonance spectroscopy
HRV3C	human rhinovirus 3C protease

IMAC	immobilised metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
k_{cat}	catalytic turnover number
kDa	kilodalton
K_i	inhibition constant
K_{ic}	competitive inhibition constant
K_{inc}	non-competitive inhibition constant
K_{iuc}	un-competitive inhibition constant
K_M	Michaelis-Menten constant
k_{off}	rate constant for dissociation
k_{on}	rate constant for association
L	litre
*lac*UV5	isopropyl β-D-1-thiogalactopyranoside-inducible promoter
LB	luria bertani
LDAO	n-dodecyl-N,N-dimethylamine-N-oxide
LDH	lactate dehydrogenase
LeuT	sodium leucine symporter
M	molar
mg	milligram
Mhp1	sodium hydantoin symporter
min	minute
mL	millilitre
mm	millimetre
mM	millimolar
MNG	2,2-didecylpropane-1,3-bis-β-D-maltopyranoside
mol	moles
MRE	mean residue ellipticity
MRSA	methicillin-resistant *Staphylococcus aureus*
msmK	shared ATPase domain
NAD+	nicotinamide adenine dinucleotide
NADH	nicotinamide adenine dinucleotide, reduced form
NADP+	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide, reduced form
NDP	nucleoside diphosphate

NanT	sugar proton symporter
NG	n-octyl-β-D-glucopyranoside
Ni-NTA	nickel nitrilotriacetic acid
nm	nanometre
OD_{600}	optical density at 600 nm
$P(r)$	real space distance distribution function
PBS	phosphate buffered saline
PCR	polymerase chain reaction
r	radius
R_{factor}	residual factor
R_{free}	free R_{factor}
RFU	relative fluorescent units
R_g	radius of gyration
*rha*BAD	L-rhamnose promoter
rmsd	root-mean-square deviation
RNA	ribonucleic acid
ROK	repressor, open reading frame, kinase
rpm	revolutions per minute
S	sedimentation coefficient, Svedberg
s	second
$s_{20,w}$	standardised sedimentation coefficient relative to water at 20 °C
[S]	substrate concentration
SatA	periplasmic substrate binding protein A
SatB/SatC	integral membrane permease domain
SatC	periplasmic substrate binding protein C
SatD	adenosine triphosphatase domain
SAXS	small angle X-ray scattering
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGLT	sodium galactose transporter
SOC	catabolite repression super optimal broth
SSS	sodium solute symporter
T7*lac*	T7-ribonucleic acid polymerase promoter
TAE	tris-acetate-ethylenediaminetetraacetic acid

TB	terrific broth
T_M	melting temperature
TRAP	tripartite adenosine triphosphate independent periplasmic transporter
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
UDP	uridine diphosphate
UV	ultraviolet
V_0	initial velocity
V_M ($Å^3$ Da^{-1})	wilson B value
v/v	volume to volume
w/v	weight to volume
Å	angstrom
°	degree
°C	degree celcius
µg	microgram
µL	microlitre
µM	micromolar
%	percent
π	pi
σ	sigma
θ	theta
$(\beta/\alpha)_8$-barrel	triosephosphate isomerase barrel
λ	wavelength

Chapter One

Introduction

The following is a general introduction to the field of sialobiology in the context of bacterial sialic acid metabolism. Subsequent chapters include more detailed, specific introductions to the chapter topics, as these are intended to be independent bodies of work.

1.1 Sialic acid

Sialic acid is the designation given to a large family comprising over 50 structurally distinct nine carbon amino sugars. All are derived from the most common compound 2-keto-3-deoxy-5-acetamido-D-glycero-D-galacto-nonulosonic acid, more commonly known as *N*-acetylneuraminic acid (Figure 1.1) (Vimr *et al.*, 2004; Almagro-Moreno & Boyd, 2009a). Derivatives of this molecule carry various substituents at the amino or hydroxyl groups (Varki, 1992; Schauer, 2000; Angata & Varki, 2002). Interestingly, *N*-acetylneuraminic acid is the only sialic acid that is ubiquitous in all organisms (Schauer, 2004). Throughout this thesis, the term sialic acid refers to *N*-acetylneuraminic acid.

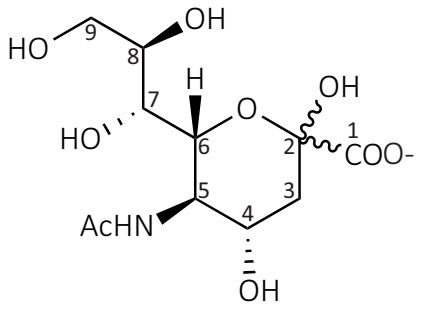


Figure 1.1. Chemical structure of *N*-acetylneuraminic acid. Carbons are numbered one through nine.

1.2 Sialic acid utilisation by bacterial pathogens

Eukaryotic cell surfaces are decorated with a complex array of glycoconjugates. Negatively charged sialic acids are found attached to the terminal sugar positions of these cell surface glycoconjugates where they mediate a diverse array of cellular interactions, recognition and adhesion (Varki, 1993; Schauer, 2000; Vimr *et al.*, 2004). In the human respiratory and gastrointestinal tract, sialic acid coated glycoconjugates are highly abundant (Vimr, 2013) and glucose tends to be limited in supply (Jeong *et al.*, 2009). Thus, bacterial pathogens that colonise these environments, along with commensal competitors, have evolved mechanisms by which they can utilise host-derived sialic acid (Vimr *et al.*, 2004; Severi *et al.*, 2007; Almagro-Moreno & Boyd, 2009a).

Although the concentration of sialic acid present in human serum is relatively high at approximately 1.6 to 2.2 mM/L, it is bound to glycoconjugates under normal physiological conditions (Sillanaukee *et al.*, 1999). By definition, these sialoglycoconjugates are unavailable for utilisation by bacterial pathogens. In order to make sialic acid available, it needs to be released by a neuraminidase, which hydrolyses the linkage between the sialic acid molecule and the sub-terminal sugar of the glycoconjugate (Vimr, 1994; Vimr, 2013). Neuraminidases are produced either endogenously by the host in response to inflammation (Sohanpal *et al.*, 2004; Sohanpal *et al.*, 2007) or exogenously by neuraminidase expressing bacteria present in the heavily sialylated niche (Shakhnovich *et al.*, 2002; Vimr, 2013).

Once available, bacterial pathogens can utilise host-derived sialic acid in multiple ways. This includes a strategy known as molecular mimicry, where bacterial pathogens coat their outer surfaces in sialic acid to circumvent the host's innate immune response (Bouchet *et al.*, 2003; Vimr *et al.*, 2004; Severi *et al.*, 2007). But most importantly, sialic acid can be degraded by some bacterial pathogens into a source of carbon, nitrogen and energy (Vimr & Troy, 1985; Olson *et al.*, 2013). Given that glucose is limited in the human respiratory or gastrointestinal tract, the utilisation of host-derived sialic acid as an alternate nutrient source provides bacterial pathogens with a selective advantage in the human host (Martinez *et al.*, 1995; Chang *et al.*, 2004; Vimr *et al.*, 2004; Severi *et al.*, 2005; Steenbergen *et al.*, 2005; Almagro-Moreno & Boyd, 2009b; Jeong *et al.*, 2009). The utilisation of sialic acid by bacterial pathogens is described in detail below.

1.3 Bacterial pathogens evade the host immune response by molecular mimicry

Some bacterial pathogens have evolved the capacity to incorporate sialic acid into cell surface macromolecules known as lipopolysaccharides and lipooligosaccharides or to produce capsules containing homopolymers of sialic acid (Vimr *et al.*, 2004). This allows for the bacterial pathogen to circumvent the host immune response through molecular mimicry. The evolutionary advantage of such mimicry is so pronounced that at least four pathways of cell surface sialylation have evolved (Vimr & Lichtensteiger, 2002). These include precursor scavenging, *de novo* biosynthesis, donor scavenging and *trans*-neuraminidase activity (Figure 1.2).

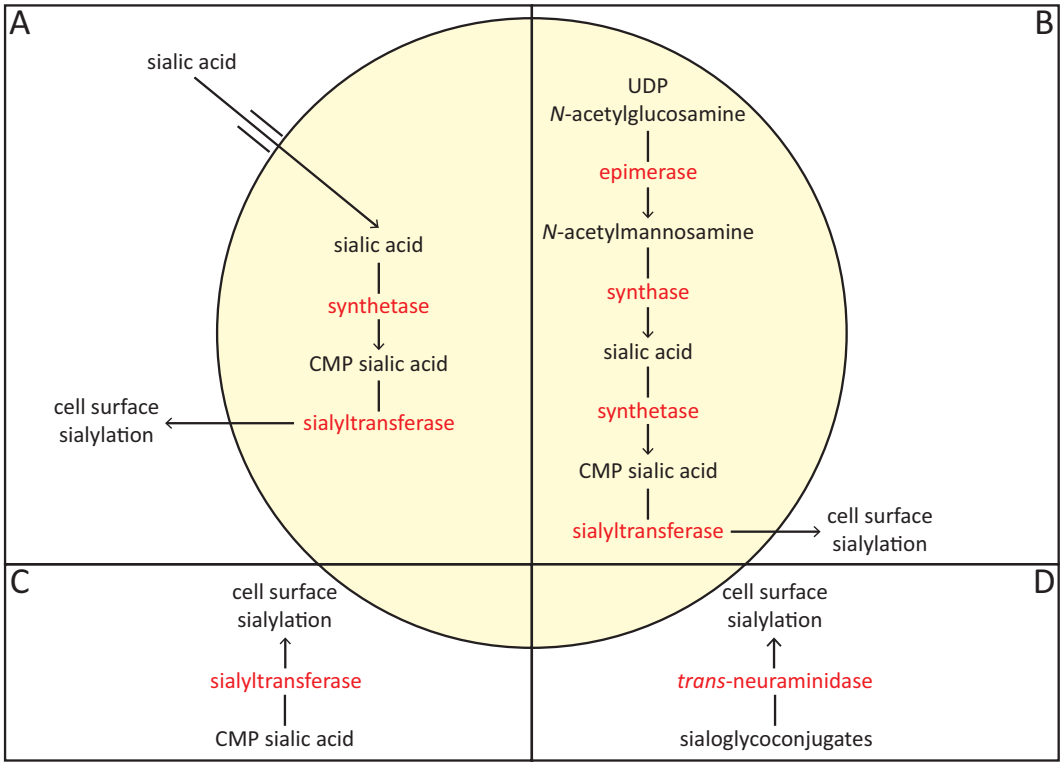


Figure 1.2. The pathways of bacterial cell surface sialylation. Four pathways have evolved that allow sialic acid to be incorporated into cell surface macromolecules. **(A)** The precursor scavenging pathway. **(B)** The *de novo* biosynthesis pathway. **(C)** The donor scavenging pathway. **(D)** The *trans*-neuraminidase pathway.

1.3.1 Precursor scavenging

A recently discovered method of cell surface sialylation known as precursor scavenging (Figure 1.2 A) has been revealed in *Haemophilus influenzae*. This organism scavenges host-derived sialic acid by importing it into its cell. Once imported, *H. influenzae* activates the sialic acid by converting it into cytidine monophosphate (CMP) sialic acid, which is incorporated into appropriate acceptors by a sialyltransferase, with the resultant sialoconjugate exported to the cell surface (Vimr & Lichtensteiger, 2002). Appropriate acceptors include lipopolysaccharides, lipooligosaccharides or capsules containing homopolymers of sialic acid. Alternatively, the imported sialic acid can be degraded into a carbon, nitrogen and energy source (Section 1.5) (Almagro-Moreno & Boyd, 2009a). Thus, *H. influenzae* must make an important metabolic decision between cell surface sialylation or sialic acid degradation so that a fine balance is maintained between the need to evade the host's immune response and any nutritional requirements (Vimr & Lichtensteiger, 2002).

1.3.2 *De novo* biosynthesis

De novo biosynthesis of sialic acid (Figure 1.2 B) is used by a number of bacteria including *Neisseria meningitides* and *Escherichia coli* (Vimr *et al.*, 2004). This route involves synthesising sialic acid from uridine diphosphate (UDP) *N*-acetylglucosamine, a simple metabolite produced by most cells. Once synthesised, sialic acid is converted into its activated form, CMP sialic acid, which is incorporated into appropriate acceptors by a sialyltransferase, with the resultant sialoconjugate exported to the cell surface (Vimr & Lichtensteiger, 2002).

1.3.3 Donor scavenging

The donor scavenging pathway of cell surface sialylation (Figure 1.2 C) has been discovered in *Neisseria gonorrhoeae*, which enables a considerably truncated sialylation pathway. Like *de novo* biosynthesis, this mechanism also involves the addition of CMP sialic acid to appropriate acceptors. But instead of synthesising it from simple metabolites endogenously, CMP sialic acid is scavenged directly from the exogenous environment and

incorporated into appropriate acceptors by a sialyltransferase, with the resultant sialoconjugate exported to the cell surface (Shell *et al.*, 2002).

1.3.4 *Trans*-neuraminidase activity

As described in Section 1.2, neuraminidases can hydrolytically remove sialic acid from host glycoconjugates, making it available for utilisation by bacterial pathogens (Vimr, 1994; Vimr, 2013). A *trans*-neuraminidase enzyme for cell surface sialylation (Figure 1.2 D) has been discovered in *Trypanosoma cruzi* with both hydrolytic and sialyltransferase activity. The hydrolytic activity is responsible for cleaving sialic acid from host glycoconjugates, while the sialyltransferase activity is responsible for directly transferring it to the cell surface of *T. cruzi* (Previato *et al.*, 1985; Vimr & Lichtensteiger, 2002).

1.4 Importation of sialic acid into bacterial pathogens

For bacterial pathogens to utilise the precursor scavenging mechanism of cell surface sialylation or degrade host-derived sialic acid into a carbon, nitrogen and energy source (Section 1.5), sialic acid must be transported across the cytoplasmic membrane and into the bacterial cell. For gram negative bacterial pathogens, sialic acid must first cross the outer cell membrane into the periplasmic space before it can be imported into the cell. Sialic acid transport across the outer cell membrane of bacterial pathogens is not well understood (Severi *et al.*, 2007). However, a sialic acid inducible outer membrane porin known as NanC has been well characterised in *E. coli* (Condemine *et al.*, 2005; Wirth *et al.*, 2009).

Bacterial pathogens have evolved multiple mechanisms of sialic acid transport across the cytoplasmic membrane, which vary between species (Almagro-Moreno & Boyd, 2009a). To date, four different types of transporter have been recognised, including: 1) an adenosine triphosphate (ATP) binding cassette (ABC) transporter (Post *et al.*, 2005); 2) a tripartite ATP independent periplasmic transporter (TRAP) (Allen *et al.*, 2005; Severi *et al.*, 2005; Chowdhury *et al.*, 2012; Mulligan *et al.*, 2012); 3) a single component sugar proton symporter (NanT) (Vimr & Troy, 1985; Martinez *et al.*, 1995); and 4) a single component sodium solute symporter (SSS) (Severi *et al.*, 2010). Each type of sialic acid transporter is described below and depicted in Figure 1.3.

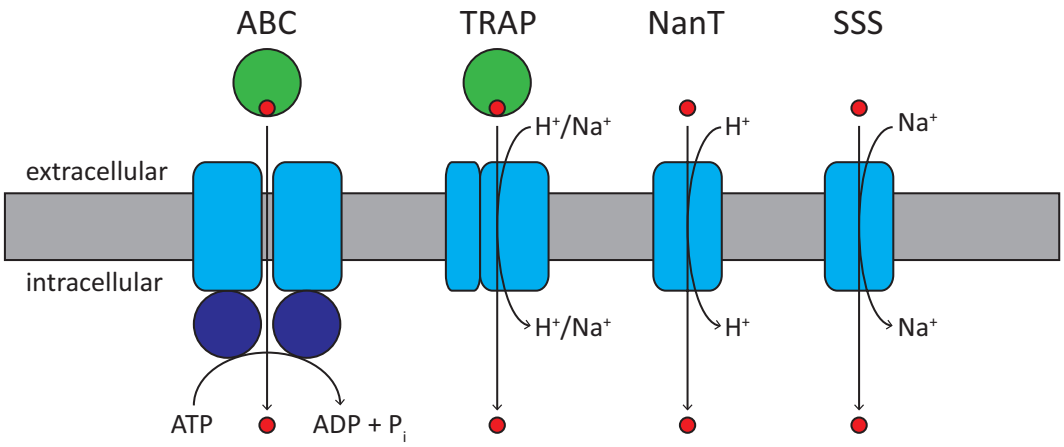


Figure 1.3. Sialic acid transporter types. In ABC and TRAP transporter systems, a periplasmic or cell surface associated substrate binding protein (green circles) interacts with two membrane protein domains (blue rectangles) to transport the sialic acid (red circles) across the membrane. ABC transporters are primary transporters that couple solute translocation with the hydrolysis of ATP by the ATPase domains harbouring the Walker motifs (dark blue circles). TRAP, NanT and SSS transporters are secondary transporters that couple solute translocation with an electrochemical gradient. It is unclear whether TRAP sialic acid transporters are proton or sodium ion dependent; but NanT sialic acid transporters have been shown to be proton dependent (Mulligan *et al.*, 2012), while SSS sialic acid transporters have been shown to be sodium ion dependent (Severi *et al.*, 2010). The stoichiometry of ions required for transport is not well understood.

1.4.1 ATP binding cassette transporters for sialic acid transport

The ABC superfamily of transporters is one of the largest families of transporters with representatives in all phyla (Jones & George, 2004). Containing uptake and efflux systems, this superfamily is made up of dozens of families that vary in their substrate specificity (Saurin *et al.*, 1999). ABC transporters are known as primary transporters because they utilise the energy generated from ATP hydrolysis to transport solutes across a membrane (Higgins, 1992). They transport a variety of substrates across both extracellular and intracellular membranes, including ions, sugars, amino acids, peptides, polysaccharides and even proteins.

All ABC transporters are composed of four domains, two transmembrane domains that form the translocation pathway and two cytoplasmic nucleotide binding domains that

hydrolyse ATP (Higgins, 1992). Prokaryotic ABC transporters contain an additional periplasmic or cell surface associated binding protein that bind substrates with high affinity and deliver them to the transmembrane domains. Since the first high resolution structure of the *E. coli* vitamin B_{12} ABC transporter (Locher *et al.*, 2002), crystal structures of a handful of others have been solved (Oldham *et al.*, 2007; Hohl *et al.*, 2012; Woo *et al.*, 2012; Shintre *et al.*, 2013), some of which have been solved in complex with their substrate binding proteins (Hollenstein *et al.*, 2007; Oldham *et al.*, 2007).

The first sialic acid ABC transport system was identified and characterised in *Haemophilus ducreyi* (Post *et al.*, 2005). More recently, an ABC transporter system for sialic acid has been identified in *Streptococcus pneumoniae* (Marion *et al.*, 2011a; Marion *et al.*, 2011b). Although the import of sialic acid by ABC transporters is not well characterised in other organisms, *Streptococcus agalactiae*, *Streptococcus gordonii, Streptococcus pyogenes* and *Streptococcus sanguinis* are predicted to utilise this type of sialic acid transport system (Almagro-Moreno & Boyd, 2009a). The components that make up the *H. ducreyi* ABC transporter include a periplasmic substrate binding protein (SatA), integral membrane permease domains (SatB/SatC) and an ATPase domain (SatD). *S. pneumoniae* has a similar structural organisation, except that the ATPase domain is a shared ATPase (msmK) that is responsible for energising multiple carbohydrate transporters (Marion *et al.*, 2011a).

1.4.2 Tripartite ATP independent periplasmic transporters for sialic acid transport

TRAP transporters constitute a large family of specific solute transporters, all of which transport organic acids. Solutes transported include a range of C4-dicarboxylates, α-keto acids, aromatic substrates and amino acids (Mulligan *et al.*, 2011). All known TRAP transporters are secondary transporters that use an electrochemical gradient to facilitate solute transport (Kelly & Thomas, 2001). Secondary transporters couple the movement of an ion down an electrochemical gradient with the movement of another ion or molecule against a concentration and/or electrochemical gradient. Unlike primary transporters, there is no direct coupling of ATP to allow movement of an ion or molecule across the membrane. By definition, secondary transporters can operate in both directions, depending

on the direction and magnitude of the concentration and/or electrochemical gradient (Poolman & Konings, 1993; Severi *et al.*, 2010).

TRAP transporters are composed of three protein domains; these include the substrate binding protein, a small membrane spanning domain and a large membrane spanning domain. The purpose of the substrate binding protein is to bind substrate with high affinity and specificity and present it to the membrane spanning domains of the transporter for transport across the membrane (Doeven *et al.*, 2004; Mulligan *et al.*, 2009). Conveniently, it is found either free in the periplasm of gram negative bacteria or anchored to the cytoplasmic membrane in gram positive bacteria and archaea (Kelly & Thomas, 2001). For drug discovery, TRAP transporters are of particular interest because they are widespread across bacteria and archaea but are not found in eukaryotes (Kelly & Thomas, 2001).

The first C4-dicarboxylate TRAP transport system to be described was from *Rhodobacter capsulatus* (Forward *et al.*, 1997). Although TRAP transporters are not well characterised, the large membrane spanning domain is predicted to form the membrane translocation pathway through which the substrate passes (Mulligan *et al.*, 2012). Whereas the small membrane spanning domain has been shown to be essential for transport, but the function remains unknown (Forward *et al.*, 1997; Mulligan *et al.*, 2012). It is unclear whether TRAP sialic acid transporters are proton or sodium ion dependent (Mulligan *et al.*, 2012).

1.4.3 Sugar proton symporters for sialic acid transport

The first sialic acid transporter to be discovered was NanT from *E. coli* (Vimr & Troy, 1985). NanT is a secondary transporter belonging to the major facilitator superfamily (Martinez *et al.*, 1995; Vimr *et al.*, 2004). NanT is a single component system that is predicted to be a sugar proton symporter, specifically co-transporting sialic acid with a proton into the cell (Vimr *et al.*, 2004; Severi *et al.*, 2010). The stoichiometry of protons required for the transport of sialic acid by NanT is yet to be confirmed. Unfortunately, our understanding of NanT at the molecular level is limited (Martinez *et al.*, 1995). In Chapter Four of this thesis, I will attempt to address this by investigating a sialic acid transporter of this type from *Yersinia pestis*, a gram negative bacterial pathogen that can infect humans and animals causing the Bubonic plague (Almagro-Moreno & Boyd, 2009a).

1.4.4 Sodium solute symporters for sialic acid transport

More recently, a novel type of sialic acid transporter has been discovered from *Salmonella enterica* (Severi *et al.*, 2010), belonging to the sodium solute symporter (SSS) family of secondary transporters that co-transport sodium ions with sugars, amino acids, inorganic ions or vitamins (Wright *et al.*, 2004). Like NanT, the SSS transporters are also single component systems. The *S. enterica* sialic acid transporter was established as a typical member of the SSS family because its functionality was dependent upon the presence of sodium ions (Severi *et al.*, 2010). The genes encoding the SSS sialic acid transporters are widespread among both gram positive and gram negative species of bacteria, whereas other types of sialic acid transporters do not appear to be as widespread (Severi *et al.*, 2010). An SSS sialic acid transporter from *S. aureus* and *Proteus mirabilis* will be explored in more detail in Chapter Four.

1.4.5 Sialic acid is transported using a specific transporter

A functional sialic acid transporter has proven to be essential for the uptake of sialic acid in a range of human bacterial pathogens, including *E. coli*, *H. influenzae*, *S. aureus* and *V. cholerae* (Vimr & Troy, 1985; Vimr *et al.*, 2004; Severi *et al.*, 2005; Severi *et al.*, 2010; Mulligan *et al.*, 2012; Olson *et al.*, 2013). This means that these organisms possess one specific type of transporter that is the sole route for sialic acid uptake. Thus, developing novel antibiotic drugs that target these transporters is a plausible mechanism for blocking pathogen colonisation. Indeed, inhibition of sialic acid transporters has lead to bacterial non-virulence in animal models (Fuller *et al.*, 2000; Chang *et al.*, 2004; Severi *et al.*, 2007). However, to date there are no structural data to explain how sialic acid transporters mediate transport across the plasma membrane or how specificity is achieved. Understanding this will be critical for the design of future antibiotics.

1.5 Sialic acid degradation by bacterial pathogens

Following the transport of host-derived sialic acid into the cell, some bacterial pathogens will utilise it by degrading it into a carbon, nitrogen and energy source (Vimr & Troy, 1985; Olson *et al.*, 2013). The genes required for the sequestration and subsequent degradation of host-derived sialic acid are known as the '*nan nag* cluster,' and are confined to pathogenic species of bacteria and mammalian commensals (Almagro-Moreno & Boyd, 2009a). The *nan nag* cluster encodes the transporter responsible for importing sialic acid into the bacterial cell and five catalytic enzymes that successively degrade it into fructose-6-phosphate, a key metabolite for glycolysis.

The sialic acid degradation pathway is depicted in Figure 1.4 and each of the catalytic enzymes will be described in more detail below. Briefly, the degradation of sialic acid is initiated by *N*-acetylneuraminate lyase, an enzyme that catalyses the retro-aldol cleavage of open chain sialic acid to form *N*-acetylmannosamine and pyruvate (Izard *et al.*, 1994; Barbosa *et al.*, 2000). *N*-Acetylmannosamine kinase then uses ATP to phosphorylate *N*-acetylmannosamine, generating *N*-acetylmannosamine-6-phosphate and adenosine diphosphate (ADP) (North *et al.*, 2014b). Thereafter, *N*-acetylmannosamine-6-phosphate 2-epimerase converts *N*-acetylmannosamine-6-phosphate into *N*-acetylglucosamine-6-phosphate *via* an epimerisation reaction (North *et al.*, 2014a). Next, *N*-acetylglucosamine-6-phosphate deacetylase removes the acetyl group from *N*-acetylglucosamine-6-phosphate and yields glucosamine-6-phosphate. Glucosamine-6-phosphate deaminase then catalyses the isomerisation and deamination of glucosamine-6-phosphate into fructose-6-phosphate (Almagro-Moreno & Boyd, 2009a).

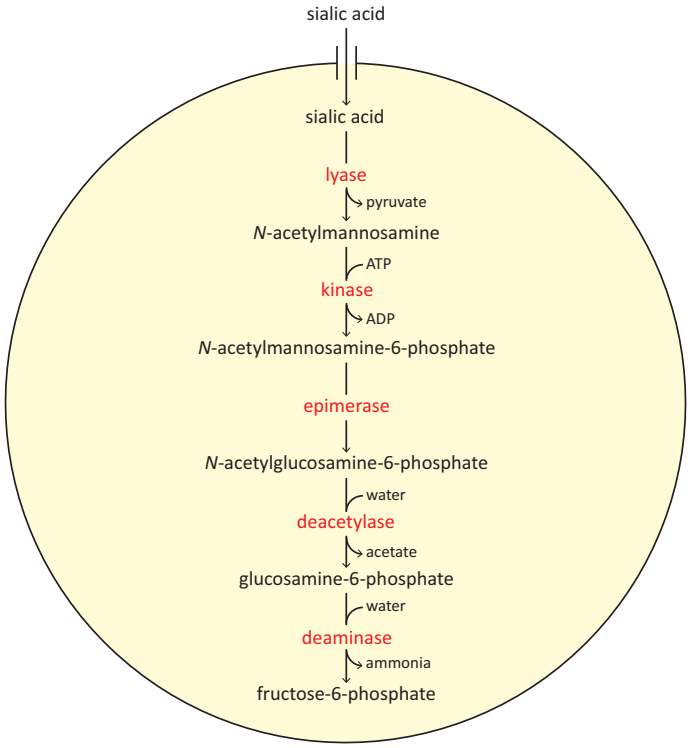


Figure 1.4. The sialic acid degradation pathway. Following the import of sialic acid into the bacterial cell, it is successively degraded into fructose-6-phosphate by *N*-acetylneuraminate lyase, *N*-acetylmannosamine-6-phosphate, *N*-acetylmannosamine-6-phosphate 2-epimerase, *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase.

1.5.1 *N*-Acetylneuraminate lyase

N-Acetylneuraminate lyase (which will be discussed in more detail in Chapter Two) catalyses the retro-aldol cleavage of open chain sialic acid to form *N*-acetylmannosamine and pyruvate (Figure 1.5) *via* a Schiff base intermediate (Izard *et al.*, 1994; Lawrence *et al.*, 1997; Barbosa *et al.*, 2000). This reaction represents the first and the committed step of the degradation pathway. Structurally, *N*-acetylneuraminate lyase is classified within the *N*-acetylneuraminate lyase sub-family of triosephosphate isomerase barrel [$(\beta/\alpha)_8$-barrel] enzymes. Members of this sub-family identified thus far include the archetype *N*-acetylneuraminate lyase, 4-hydroxy-tetrahydrodipicolinic acid synthase [formally known as dihydrodipicolinate synthase (DHDPS)] (Mirwaldt *et al.*, 1995), D-5-keto-4-deoxyglucarate dehydratase (Jeffcoat *et al.*, 1969b; Jeffcoat *et al.*, 1969a), 2-keto-3-deoxygluconate aldolase (Buchanan *et al.*, 1999), *trans*-*o*-hydroxybenzylidenepyruvate

hydratase-aldolase (Eaton, 1994) and *trans*-2'-carboxybenzalpyruvate hydratase-aldolase (Iwabuchi & Harayama, 1998). All of which share a common structural framework, but catalyse reactions in separate biochemical pathways (Lawrence *et al.*, 1997).

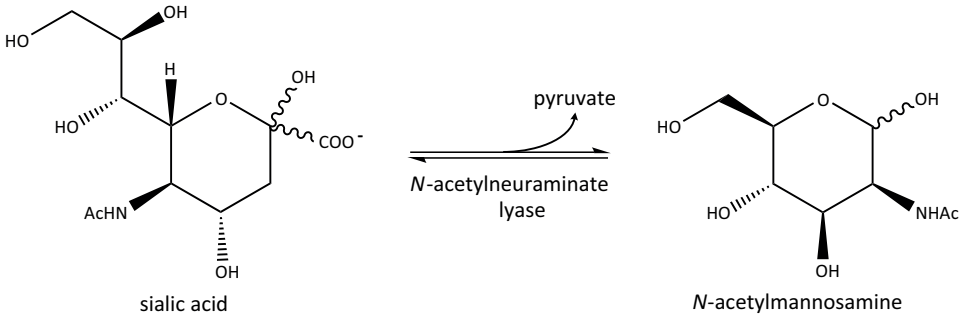


Figure 1.5. The reaction catalysed by *N*-acetylneuraminate lyase. *N*-Acetylneuraminate lyase catalyses the retro-aldol cleavage of open chain sialic acid to form *N*-acetylmannosamine and pyruvate.

The structure of *N*-acetylneuraminate lyase has been well characterised in various organisms including, *E. coli* (Izard *et al.*, 1994; Lawrence *et al.*, 1997; Joerger *et al.*, 2003; Campeotto *et al.*, 2009; Campeotto *et al.*, 2010; Daniels *et al.*, 2014), *H. influenzae* (Barbosa *et al.*, 2000), *Pasteurella multocida* (Huynh *et al.*, 2013) and *S. aureus* (Timms *et al.*, 2013). Further analysis of *N*-acetylneuraminate lyase in complex with substrate analogues has revealed the mechanism of substrate binding within the active site (Lawrence *et al.*, 1997; Barbosa *et al.*, 2000; Huynh *et al.*, 2013; Daniels *et al.*, 2014). As with all members of the *N*-acetylneuraminate lyase sub-family of $(\beta/\alpha)_8$-barrel enzymes, catalysis is proposed to involve the formation of a Schiff base between the amine of a highly conserved lysine residue and the C2 carbon of the α-keto acid moiety on the substrate (Izard *et al.*, 1994; Blickling *et al.*, 1997; Lawrence *et al.*, 1997). In recent years, *N*-acetylneuraminate lyase has received considerable attention from both mechanistic and structural viewpoints and has been recognised as a viable antibiotic drug target (Severi *et al.*, 2007; von Itzstein, 2007).

1.5.2 *N*-Acetylmannosamine kinase

Following the reaction catalysed by *N*-acetylneuraminate lyase, *N*-acetylmannosamine kinase transfers a phosphate group from ATP to the C6 position of *N*-acetylmannosamine,

generating *N*-acetylmannosamine-6-phosphate and ADP (Figure 1.6). *N*-Acetylmannosamine kinase is a phosphotransferase belonging to the repressor, open reading frame, kinase (ROK) superfamily. ROK kinases usually contain a conserved N-terminal ATP binding motif and a zinc binding motif (Holmes *et al.*, 1993; Larion *et al.*, 2007). Zinc is proposed to play a key structural role in the active site of many ROK kinases (Larion *et al.*, 2007). However, this zinc binding motif is not obvious in *N*-acetylmannosamine kinase from *S. aureus* and *Streptococcus mitis* (North *et al.*, 2014b). The lack of a zinc binding motif in these organisms suggests that the geometry and stability of the active site may be affected in some other way. This difference could be explored as a target for inhibitory molecules.

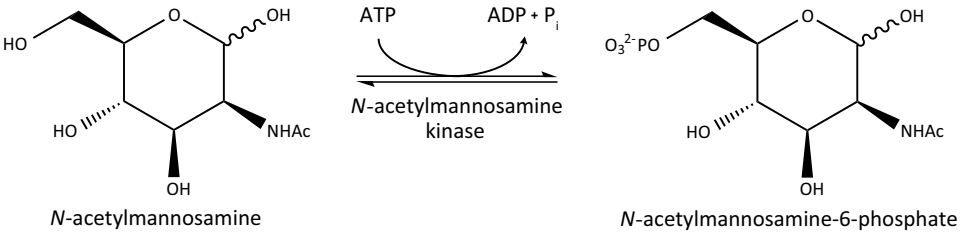


Figure 1.6. The reaction catalysed by *N*-acetylmannosamine kinase. *N*-Acetylmannosamine kinase transfers a phosphate group from ATP to the C6 position of *N*-acetylmannosamine, generating *N*-acetylmannosamine-6-phosphate and ADP.

There is a lack of literature published on the bacterial *N*-acetylmannosamine kinase, but much is known about its bifunctional eukaryotic homologue UDP *N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase. In mammals, this enzyme is proposed to be involved in sialic acid biosynthesis (Hinderlich *et al.*, 1997). Crystal structures of the *Homo sapiens N*-acetylmannosamine kinase domain of UDP *N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase have been solved (Martinez *et al.*, 2012). Yet, only one bacterial crystal structure of *N*-acetylmannosamine kinase has been determined from *E. coli* (New York Structural Genomics Research Centre, unpublished work).

1.5.3 *N*-Acetylmannosamine-6-phosphate 2-epimerase

N-Acetylmannosamine-6-phosphate 2-epimerase (which will be investigated in more detail in Chapter Three) converts *N*-acetylmannosamine-6-phosphate into *N*-acetylglucosamine-

6-phosphate (Figure 1.7). *N*-acetylmannosamine-6-phosphate 2-epimerase and two other amino sugar 2-epimerase enzymes have been implicated in pathways that utilise sialic acid in both eukaryotes and prokaryotes, all of which use *N*-acetylmannosamine as the essential sugar precursor (Ferrero *et al.*, 2007). In eukaryotes, *N*-acetylglucosamine 2-epimerase is responsible for the degradation of sialic acid, whereas the previously discussed bifunctional UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase is involved in the biosynthesis of sialic acid. An evolutionarily and mechanistically related UDP *N*-acetylglucosamine 2-epimerase has been implicated in the biosynthesis of sialic acid in prokaryotes (Tanner, 2005), but it lacks the bifunctional kinase activity.

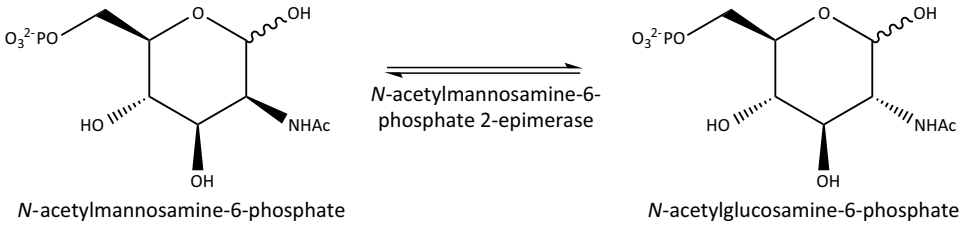


Figure 1.7. The reaction catalysed by *N*-acetylmannosamine-6-phosphate 2-epimerase. *N*-Acetylmannosamine-6-phosphate 2-epimerase epimerises *N*-acetylmannosamine-6-phosphate into *N*-acetylglucosamine-6-phosphate.

Crystal structures of *N*-acetylmannosamine-6-phosphate 2-epimerase from methicillin-resistant *S. aureus* (MRSA) (Midwest Centre for Structural Genomics, unpublished work), *S. enterica* (Centre for Structural Genomics of Infectious Diseases, unpublished work) and *S. pyogenes* (Midwest Centre for Structural Genomics, unpublished work) have been solved. However, these structures have not yet been published in the literature. Recently however, the structure of *Clostridium perfringens N*-acetylmannosamine-6-phosphate 2-epimerase has been solved and published (Pelissier *et al.*, 2014).

1.5.4 *N*-Acetylglucosamine-6-phosphate deacetylase

Following the reaction catalysed by *N*-acetylmannosamine-6-phosphate 2-epimerase, *N*-acetylglucosamine-6-phosphate deacetylase catalyses the deacetylation of *N*-acetylglucosamine-6-phosphate with water to form glucosamine-6-phosphate and acetate (Figure 1.8). Crystal structures of *N*-acetylglucosamine-6-phosphate deacetylase have been

reported from *E. coli* in a metal free state (Ferreira *et al.*, 2006), *E. coli* in a metal bound state (Hall *et al.*, 2007), *Bacillus subtilis* in a metal bound state (Vincent *et al.*, 2004) and *Thermotoga maritima* in a metal bound state (Joint Centre for Structural Genomics, unpublished work). These structures show that the *N*-acetylglucosamine-6-phosphate deacetylases belong to the metal dependent amidohydrolase superfamily (Vincent *et al.*, 2004). This family of enzymes can bind one or more divalent metal ions in the active site, including Fe^{2+}, Zn^{2+}, Co^{2+} and Cu^{2+} (Ferreira *et al.*, 2006). A comparison of these crystal structures reveals different active site motifs for the complexation of divalent metal ions. The structure for *B. subtilis* *N*-acetylglucosamine-6-phosphate deacetylase contains a binuclear Fe^{2+} centre, whereas the *T. maritima* structure *N*-acetylglucosamine-6-phosphate deacetylase has a mononuclear Fe^{2+} centre (Vincent *et al.*, 2004). The structure of the *E. coli* enzyme in a metal bound state has a mononuclear Fe^{2+} centre (Hall *et al.*, 2007), yet this enzyme has also been shown to bind other metal ions (Ferreira *et al.*, 2006).

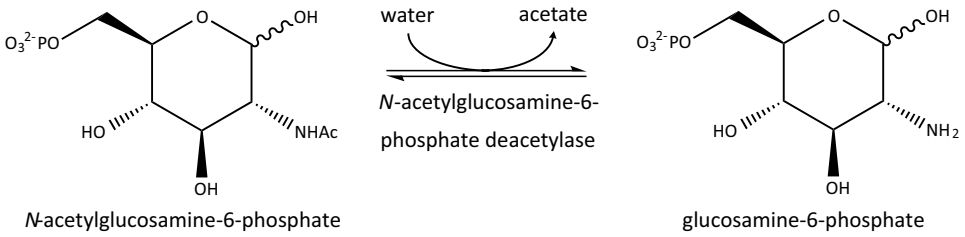


Figure 1.8. The reaction catalysed by *N*-acetylglucosamine-6-phosphate deacetylase. *N*-Acetylglucosamine-6-phosphate deacetylase catalyses the deacetylation of *N*-acetylglucosamine-6-phosphate with water to form glucosamine-6-phosphate and acetate.

1.5.5 Glucosamine-6-phosphate deaminase

The final enzyme involved in the degradation of sialic acid is glucosamine-6-phosphate deaminase. This enzyme catalyses the isomerisation and deamination of glucosamine-6-phosphate with water into fructose-6-phosphate and ammonia (Figure 1.9). Fructose-6-phosphate is an essential metabolite that is used in glycolysis (Almagro-Moreno & Boyd, 2009a). Glucosamine-6-phosphate deaminase belongs to an aldose ketose isomerase class of proteins, all of which catalyse the removal of one hydrogen from the C2 position of the aldose and add back a hydrogen to the C1 position of the ketose through a *cis*-enediol intermediate (Oliva *et al.*, 1995). This enzyme has been identified in several species of

animal, fungus and bacteria, where it is involved in sialic acid degradation for nutritional purposes or amino sugar synthesis for peptidoglycan formation (Liu *et al.*, 2008). The equilibrium of the reaction favours the degradation reaction (Calcagno *et al.*, 1984).

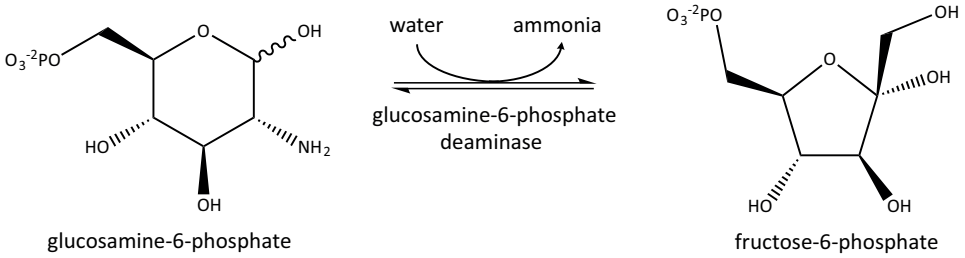


Figure 1.9. The reaction catalysed by glucosamine-6-phosphate deaminase. Glucosamine-6-phosphate deaminase catalyses the isomerisation and deamination of glucosamine-6-phosphate with water into fructose-6-phosphate and ammonia.

Crystal structures of glucosamine-6-phosphate deaminase have been published from a range of bacterial species including, *B. subtilis* (Vincent *et al.*, 2005), *E. coli* (Oliva *et al.*, 1995; Horjales *et al.*, 1999; Rudino-Pinera *et al.*, 2002) and *Streptococcus mutans* (Liu *et al.*, 2008). The crystal structure of glucosamine-6-phosphate deaminase from *E. coli* shows that the enzyme is hexameric and allosterically activated by *N*-acetylglucosamine-6-phosphate (Horjales *et al.*, 1999). In contrast, glucosamine-6-phosphate deaminase from *B. subtilis* and *S. mutans* is monomeric and not allosterically regulated (Vincent *et al.*, 2005; Liu *et al.*, 2008). It has been proposed that through evolution, glucosamine-6-phosphate deaminase from gram positive bacteria has diverged into a monomeric family, causing it to lose the property of allosteric regulation (Liu *et al.*, 2008).

1.6 Regulation of the *nan nag* cluster

Interestingly, sialic acid degradation does not appear to be under strict allosteric control, apart from the activation of glucosamine-6-phosphate deaminase by *N*-acetylglucosamine-6-phosphate. Instead, expression of the enzymes involved in the transport and degradation of sialic acid is controlled by a transcriptional regulator, known as the nanRepressor (Kalivoda *et al.*, 2003; Johnston *et al.*, 2007; Hwang *et al.*, 2013; Kalivoda *et al.*, 2013; Olson *et al.*, 2013). The nanRepressor belongs to the RpiR family of transcriptional regulators and contains a helix-turn-helix motif and a sugar isomerase domain (Jaeger & Mayer, 2008). An exception to the RpiR family of transcriptional regulators is the nanRepressor found in *E. coli,* which belongs to the GntR superfamily. Despite this exception, all nanRepressor proteins are responsible for regulating the expression of the sialic acid degradation operon, thus they are also known as 'sialoregulators'.

Helix-turn-helix motifs are often found in proteins that are known to bind deoxyribonucleic acid (DNA) and regulate gene expression. The motif is characterised by two α-helices that make intimate contacts with the DNA, one of which is important for sequence recognition, while the other assists in stabilising the structure (Sauer *et al.*, 1982; Brennan & Matthews, 1989). For nanRepressor proteins belonging to the RpiR superfamily, positively charged amino acid residues of the helix-turn-helix domain are speculated to interact with the *nan* promoter. For the *S. aureus* nanRepressor in particular, the positively charged amino acid residues of the helix-turn-helix domain specifically interact with the *nan* promoter at the *nan*AT and *nan*E transcripts (Olson *et al.*, 2013).

The sugar isomerase domain of the nanRepressor has a key regulatory role in the expression of the enzymes involved in sialic acid degradation. Depending on the organism, the sugar isomerase domain interacts with the sugar-like structure of sialic acid or degradation products within the pathway (Kalivoda *et al.*, 2003; Kalivoda *et al.*, 2013; Olson *et al.*, 2013). For *E. coli*, sialic acid has been implicated as the inducer of *nan* expression (Brennan, 1993; Busby & Ebright, 1999; Blanco *et al.*, 2002; Kalivoda *et al.*, 2003; Kalivoda *et al.*, 2013). The second and third substrate/product involved in the degradation of sialic acid (*N*-acetylmannosamine and *N*-acetylmannosamine-6-phosphate, respectively), act as inducers of *nan* expression in *S. pneumoniae* (Gualdi *et al.*, 2012). *N*-

acetylmannosamine-6-phosphate alone acts as the inducer for *V. vulnificus* (Kim *et al.*, 2011) and *S. aureus* (Olson *et al.*, 2013). For *H. influenzae*, glucosamine-6-phosphate is responsible for both the induction and the repression of *nan* expression (Johnston *et al.*, 2007).

To date, the only crystal structure that has been solved for a nanRepressor protein belongs to *V. vulnificus* in complex with its regulatory ligand, *N*-acetylmannosamine-6-phosphate (Hwang *et al.*, 2013). This nanRepressor been demonstrated to form a dimer where the positioning of each monomer creates a positively charged arched tunnel-like DNA binding pocket. *N*-Acetylmannosamine-6-phosphate is known to block DNA binding. This is proposed to be mediated by relocation of residues that alleviate the interaction between the nanRepresor and DNA, relieving the repressive effect and thus inducing transcription of the *nan* genes.

1.7 Antibiotic resistance in bacterial pathogens

The discovery of potent antibiotic agents was one of the greatest advances in the control of bacterial diseases in the 20th century (Cohen, 1992). In the United States of America alone, it is estimated that more than 2 million people are infected with antibiotic resistant bacteria resulting in 23 000 deaths annually (Hampton, 2013). There are no reliable estimates of the number of infection cases globally, but there is evidence that antibiotic resistance is significantly worse in developing countries (Toner *et al.*, 2015). The rise in antibiotic resistant bacteria stems from a multitude of factors including the widespread and sometimes inappropriate administration of antibiotics in both human medicine and agriculture (Lowy, 2003; Blair *et al.*, 2015). Frighteningly, this increasing resistance has resulted in a pronounced lack in the development of new antibiotics (Blair *et al.*, 2015). Thus, the search for antibiotic compounds has recently taken on a new urgency.

The global increase in antibiotic resistance is of particular concern for *S. aureus*; a gram positive bacterial pathogen with an ability to cause a diverse range of life threatening infections (Furuya & Lowy, 2006). Since Alexander Fleming's first observation in 1928 that penicillin had antibiotic properties against staphylococci (Hare, 1982), *S. aureus* has continuously evolved resistance mechanisms against novel antibiotic compounds (Furuya & Lowy, 2006). In 1959, the first semi synthetic β-lactam antibiotic named methicillin was introduced. Despite initial successes, by 1961 the first MRSA strains were emerging, which were resistant to virtually all β-lactam antibiotics and their derivatives (Jevons, 1961). As a result of this selective pressure, there has been a dramatic increase in the number of multi-drug resistant *S. aureus* infections worldwide (McDonald *et al.*, 1981; Chambers, 1988; Panlilio *et al.*, 1992; Boyce *et al.*, 2005). Previously associated with exposure in healthcare settings, this superbug now accounts for an increasing number of super adaptable, community acquired, antibiotic resistant *S. aureus* infections (David & Daum, 2010).

1.8 Sialic acid degradation is essential for pathogen colonisation and persistence

The sialic acid degradation pathway has been well documented in several bacterial pathogens that colonise heavily sialylated niches, including *E. coli*, *H. influenzae*, *S. aureus*, *V. cholerae*, *V. vulnificus* and *Y. pestis* (Chang *et al.*, 2004; Almagro-Moreno & Boyd, 2009b; Almagro-Moreno & Boyd, 2009a; Jeong *et al.*, 2009; Olson *et al.*, 2013). As described in Section 1.2, glucose is limited in heavily sialylated niches, such as the human respiratory and gastrointestinal tract (Jeong *et al.*, 2009). Thus, the capability of these organisms to utilise sialic acid as an alternate source of carbon, nitrogen and energy is important for their survival (Almagro-Moreno & Boyd, 2009b).

Various *in vitro* and *in vivo* studies demonstrate that the degradation of sialic acid is crucial for the colonisation and persistence of several human bacterial pathogens. Deletion of one or more of the genes encoded by the *nan nag* cluster rendered *E. coli*, *S. aureus*, *V. cholerae*, *V. vulnificus* and *S. aureus* incapable of growing on sialic acid as a sole carbon, nitrogen and energy source *in vitro* (Vimr & Troy, 1985; Almagro-Moreno & Boyd, 2009b; Jeong *et al.*, 2009; Vogel-Scheel *et al.*, 2010; Olson *et al.*, 2013), which suggests that this is the only pathway for sialic degradation in these bacteria. More importantly, the ability to degrade sialic acid has proven necessary for pathogen colonisation and persistence *in vivo*, with mouse models for *E. coli* (Chang *et al.*, 2004), *V. cholerae* (Almagro-Moreno & Boyd, 2009b) and *V. vulnificus* (Jeong *et al.*, 2009). In addition to the essentiality of sialic acid transporters for the uptake of sialic acid into the bacterial cell, described in Section 1.4.5, these findings demonstrate that the *nan nag* cluster constitutes a novel target for antibiotic drug design against dangerous human bacterial pathogens.

1.9 Aims and objectives of this thesis

The overall aim of this research is to understand how sialic acid is transported into the bacterial cell and subsequently degraded. Following the import of sialic acid into the bacterial pathogen, five successive enzymes degrade it into a carbon, nitrogen and energy source. Given the importance of sialic acid transport and degradation for the colonisation and persistence of clinically important human pathogens, surprisingly little structural data exists for this pathway. Enzymology and structural biology is used to gain valuable information on how this pathway functions. These data underpin and inform future drug design and expand our understanding of this important, but poorly understood, metabolic pathway.

Chapter Two investigates the structure, function and inhibition of MRSA *N*-acetylneuraminate lyase. Biophysical characterisation, kinetic analysis and a structural analysis of MRSA *N*-acetylneuraminate lyase are presented. A previously characterised molecule that has been shown to inhibit *C. perfringens N*-acetylneuraminate lyase was synthesised in collaboration with Professor Antony Fairbanks (University of Canterbury). The ability of this molecule to inhibit MRSA *N*-acetylneuraminate lyase was tested by kinetic analysis. To probe how inhibition is achieved in this organism, the structure was solved in complex with this molecule.

Chapter Three investigates the structure and catalytic mechanism of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. A sequence and structural analysis of my MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase was used to probe amino acid residues that may be important for catalysis. These residues were mutated and a multi-enzyme coupled assay that I specifically designed for *N*-acetylmannosamine-6-phosphate 2-epimerase was used to assess the importance of these residues. During the course of this thesis, a structural analysis and proton nuclear magnetic resonance spectroscopy (^{1}H NMR) based kinetic analysis of *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase was published (Pelissier *et al.*, 2014). From this data, the authors proposed that the catalytic mechanism employed by *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes is a deprotonation/reprotonation mechanism. A critical analysis of their structures suggests that this proposed mechanism is unlikely. Alternatively, I propose that catalysis

may occur *via* the formation of a Schiff base or a proton displacement mechanism mediated by the substrate.

Chapter Four explores three sialic acid transporter proteins, belonging to two distinct families, the NanT sialic acid transporter from *Y. pestis* and the SSS sialic acid transporters from *P. mirabilis* and *S. aureus*. There are no structural data available for any of the known sialic acid transporters to explain how they mediate sialic acid transport across the cytoplasmic membrane. This chapter describes the expression, purification, quaternary structure and attempts at crystallisation of these transporters, which is a significant step towards a structural and functional understanding of these inherently difficult to work with proteins. This work was completed in collaboration with Dr Rosmarie Friemann's laboratory (University of Gothenburg). Excitingly, the structure of the *P. mirabilis* SSS sialic acid transporter was solved in Dr Rosmarie Friemann's laboratory and I completed a structural analysis, which is presented in this chapter. In addition, Chapter Four also addresses a potential protein-protein interaction between the *S. aureus* SSS sialic acid transporter and MRSA *N*-acetylneuraminate lyase enzyme, which was investigated using analytical ultracentrifugation.

1.10 References

Allen, S., Zaleski, A., Johnston, J. W., Gibson, B. W. & Apicella, M. A. (2005). Novel sialic acid transporter of *Haemophilus influenzae*. *Infection and Immunity, 73*, 5291-5300.

Almagro-Moreno, S. & Boyd, E. F. (2009a). Insights into the evolution of sialic acid catabolism among bacteria. *BioMed Central Evolutionary Biology, 9*, 118-133.

Almagro-Moreno, S. & Boyd, E. F. (2009b). Sialic acid catabolism confers a competitive advantage to pathogenic *Vibrio cholerae* in the mouse intestine. *Infection and Immunity, 77*, 3807-3816.

Angata, T. & Varki, A. (2002). Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chemical Reviews, 102*, 439-469.

Barbosa, J., Smith, B., DeGori, R., Ooi, H., Marcuccio, S., Campi, E., Jackson, W., Brossmer, R., Sommer, M. & Lawrence, M. (2000). Active site modulation in the *N*-acetylneuraminate lyase sub-family as revealed by the structure of the inhibitor-complexed *Haemophilus influenzae* enzyme. *Journal of Molecular Biology, 303*, 405-421.

Blair, J. M., Webber, M. A., Baylay, A. J., Ogbolu, D. O. & Piddock, L. J. (2015). Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology, 13*, 42-51.

Blanco, A. G., Sola, M., Gomis-Ruth, F. X. & Coll, M. (2002). Tandem DNA recognition by PhoB, a two-component signal transduction transcriptional activator. *Structure, 10*, 701-713.

Blickling, S., Renner, C., Laber, B., Pohlenz, H. D., Holak, T. A. & Huber, R. (1997). Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochemistry, 36*, 24-33.

Bouchet, V., Hood, D. W., Li, J., Brisson, J. R., Randle, G. A., Martin, A., Li, Z., Goldstein, R., Schweda, E. K., Pelton, S. I., Richards, J. C. & Moxon, E. R. (2003). Host-derived sialic acid is incorporated into *Haemophilus influenzae* lipopolysaccharide and is a major virulence factor in experimental otitis media. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 8898-8903.

Boyce, J. M., Cookson, B., Christiansen, K., Hori, S., Vuopio-Varkila, J., Kocagoz, S., Oztop, A. Y., Christiansen, K., Hori, S., Vuopio-Varkila, J., Kocagoz, S., Oztop, A. Y., Vandenbroucke-Grauls, C. M. J. E., Harbarth, S. & Pittet, D. (2005). Meticillin-resistant *Staphylococcus aureus*. *Lancet Infectious Diseases, 5*, 653-663.

Brennan, R. G. (1993). The winged-helix DNA-binding motif - another helix-turn-helix takeoff. *Cell, 74*, 773-776.

Brennan, R. G. & Matthews, B. W. (1989). The helix-turn-helix DNA-binding motif. *Journal of Biological Chemistry, 264*, 1903-1906.

Buchanan, C. L., Connaris, H., Danson, M. J., Reeve, C. D. & Hough, D. W. (1999). An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. *The Biochemical Journal, 343*, 563-570.

Busby, S. & Ebright, R. H. (1999). Transcription activation by catabolite activator protein (CAP). *Journal of Molecular Biology, 293*, 199-213.

Calcagno, M., Campos, P. J., Mulliert, G. & Suastegui, J. (1984). Purification, molecular and kinetic properties of glucosamine-6-phosphate isomerase (deaminase) from *Escherichia coli*. *Biochimica et Biophysica Acta, 787*, 165-173.

Campeotto, I., Bolt, A. H., Harman, T. A., Dennis, C., Trinh, C. H., Phillips, S. E. V., Nelson, A., Pearson, A. R. & Berry, A. (2010). Structural insights into substrate specificity in variants of *N*-acetylneuraminic acid lyase produced by directed evolution. *Journal of Molecular Biology, 404*, 56-69.

Campeotto, I., Carr, S. B., Trinh, C. H., Nelson, A. S., Berry, A., Phillips, S. E. & Pearson, A. R. (2009). Structure of an *Escherichia coli N*-acetyl-D-neuraminic acid lyase mutant, E192N, in complex with pyruvate at 1.45 Å resolution. *Acta crystallographica Section F, Structural Biology and Crystallisation Communications, 65*, 1088-1090.

Chambers, H. F. (1988). Methicillin-resistant staphylococci. *Clinical and Microbiology Reviews, 1*, 173-186.

Chang, D. E., Smalley, D. J., Tucker, D. L., Leatham, M. P., Norris, W. E., Stevenson, S. J., Anderson, A. B., Grissom, J. E., Laux, D. C., Cohen, P. S. & Conway, T. (2004). Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 7427-7432.

Chowdhury, N., Norris, J., McAlister, E., Lau, S. Y., Thomas, G. H. & Boyd, E. F. (2012). The VC1777-VC1779 proteins are members of a sialic acid-specific subfamily of

TRAP transporters (SiaPQM) and constitute the sole route of sialic acid uptake in the human pathogen *Vibrio cholerae*. *Microbiology, 158*, 2158-2167.

Cohen, M. L. (1992). Epidemiology of drug resistance: implications for a post-antimicrobial era. *Science, 257*, 1050-1055.

Condemine, G., Berrier, C., Plumbridge, J. & Ghazi, A. (2005). Function and expression of an *N*-acetylneuraminic acid-inducible outer membrane channel in *Escherichia coli*. *Journal of Bacteriology, 187*, 1959-1965.

Daniels, A. D., Campeotto, I., van der Kamp, M. W., Bolt, A. H., Trinh, C. H., Phillips, S. E., Pearson, A. R., Nelson, A., Mulholland, A. J. & Berry, A. (2014). Reaction mechanism of *N*-acetylneuraminic acid lyase revealed by a combination of crystallography, QM/MM simulation, and mutagenesis. *American Chemical Society Chemical Biology, 9*, 1025-1032.

David, M. Z. & Daum, R. S. (2010). Community-associated methicillin-resistant *Staphylococcus aureus*: epidemiology and clinical consequences of an emerging epidemic. *Clinical and Microbiology Reviews, 23*, 616-687.

Doeven, M. K., Abele, R., Tampe, R. & Poolman, B. (2004). The binding specificity of OppA determines the selectivity of the oligopeptide ATP-binding cassette transporter. *Journal of Biological Chemistry, 279*, 32301-32307.

Eaton, R. W. (1994). Organisation and evolution of naphthalene catabolic pathways: sequence of the DNA encoding 2-hydroxychromene-2-carboxylate isomerase and *trans*-*o*-hydroxybenzylidenepyruvate hydratase-aldolase from the NAH7 plasmid. *Journal of Bacteriology, 176*, 7757-7762.

Ferreira, F. M., Mendoza-Hernandez, G., Castaneda-Bueno, M., Aparicio, R., Fischer, H., Calcagno, M. L. & Oliva, G. (2006). Structural analysis of *N*-acetylglucosamine-6-phosphate deacetylase apoenzyme from *Escherichia coli*. *Journal of Molecular Biology, 359*, 308-321.

Ferrero, M. A., Martinez-Blanco, H., Lopez-Velasco, F. F., Ezquerro-Saenz, C., Navasa, N., Lozano, S. & Rodriguez-Aparicio, L. B. (2007). Purification and characterisation of GlcNAc-6-P 2-epimerase from *Escherichia coli* K92. *Acta Biochimica Polonica, 54*, 387-399.

Forward, J. A., Behrendt, M. C., Wyborn, N. R., Cross, R. & Kelly, D. J. (1997). TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *Journal of Bacteriology, 179*, 5482-5493.

Fuller, T. E., Kennedy, M. J. & Lowery, D. E. (2000). Identification of *Pasteurella multocida* virulence genes in a septicemic mouse model using signature-tagged mutagenesis. *Microbial Pathogenesis, 29*, 25-38.

Furuya, E. Y. & Lowy, F. D. (2006). Antimicrobial-resistant bacteria in the community setting. *Nature Reviews Microbiology, 4*, 36-45.

Gualdi, L., Hayre, J. K., Gerlini, A., Bidossi, A., Colomba, L., Trappetti, C., Pozzi, G., Docquier, J. D., Andrew, P., Ricci, S. & Oggioni, M. R. (2012). Regulation of neuraminidase expression in *Streptococcus pneumoniae*. *BioMed Central Microbiology, 12*, 200-211.

Hall, R. S., Brown, S., Fedorov, A. A., Fedorov, E. V., Xu, C., Babbitt, P. C., Almo, S. C. & Raushel, F. M. (2007). Structural diversity within the mononuclear and binuclear active sites of *N*-acetyl-D-glucosamine-6-phosphate deacetylase. *Biochemistry, 46*, 7953-7962.

Hampton, T. (2013). Report reveals scope of US antibiotic resistance threat. *Journal of the American Medical Association, 310*, 1661-1663.

Hare, R. (1982). New light on the history of penicillin. *Medical History, 26*, 1-24.

Higgins, C. F. (1992). ABC Transporters - from microorganisms to man. *Annual Review of Cell Biology, 8*, 67-113.

Hinderlich, S., Stasche, R., Zeitler, R. & Reutter, W. (1997). A bifunctional enzyme catalyses the first two steps in *N*-acetylneuraminic acid biosynthesis of rat liver - purification and characterisation of UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase. *Journal of Biological Chemistry, 272*, 24313-24318.

Hohl, M., Briand, C., Grutter, M. G. & Seeger, M. A. (2012). Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nature Structural & Molecular Biology, 19*, 395-402.

Hollenstein, K., Frei, D. C. & Locher, K. P. (2007). Structure of an ABC transporter in complex with its binding protein. *Nature, 446*, 213-216.

Holmes, K. C., Sander, C. & Valencia, A. (1993). A new ATP-binding fold in actin, hexokinase and Hsc70. *Trends in Cell Biology, 3*, 53-59.

Horjales, E., Altamirano, M. M., Calcagno, M. L., Garratt, R. C. & Oliva, G. (1999). The allosteric transition of glucosamine-6-phosphate deaminase: the structure of the T state at 2.3 angstrom resolution. *Structure with Folding and Design, 7*, 527-537.

Huynh, N., Aye, A., Li, Y., Yu, H., Cao, H., Tiwari, V. K., Shin, D. W., Chen, X. & Fisher, A. J. (2013). Structural basis for substrate specificity and mechanism of *N*-

acetyl-D-neuraminic acid lyase from *Pasteurella multocida*. *Biochemistry, 52*, 8570-8579.

Hwang, J., Kim, B. S., Jang, S. Y., Lim, J. G., You, D. J., Jung, H. S., Oh, T. K., Lee, J. O., Choi, S. H. & Kim, M. H. (2013). Structural insights into the regulation of sialic acid catabolism by the *Vibrio vulnificus* transcriptional repressor NanR. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 2829-2837.

Iwabuchi, T. & Harayama, S. (1998). Biochemical and genetic characterisation of *trans*-2'-carboxybenzalpyruvate hydratase-aldolase from a phenanthrene-degrading *Nocardioides* strain. *Journal of Bacteriology, 180*, 945-949.

Izard, T., Lawrence, M. C., Malby, R. L., Lilley, G. G. & Colman, P. M. (1994). The three-dimensional structure of *N*-acetylneuraminate lyase from *Escherichia coli*. *Structure, 2*, 361-369.

Jaeger, T. & Mayer, C. (2008). The transcriptional factors MurR and catabolite activator protein regulate *N*-acetylmuramic acid catabolism in *Escherichia coli*. *Journal of Bacteriology, 190*, 6598-6608.

Jeffcoat, R., Hassall, H. & Dagley, S. (1969a). The metabolism of D-glucarate by *Pseudomonas acidovorans*. *The Biochemical Journal, 115*, 969-976.

Jeffcoat, R., Hassall, H. & Dagley, S. (1969b). Purification and properties of D-4-deoxy-5-oxoglucarate hydrolyase (decarboxylating). *The Biochemical Journal, 115*, 977-983.

Jeong, H. G., Oh, M. H., Kim, B. S., Lee, M. Y., Han, H. J. & Choi, S. H. (2009). The capability of catabolic utilisation of *N*-acetylneuraminic acid, a sialic acid, is essential for *Vibrio vulnificus* pathogenesis. *Infection and Immunity, 77*, 3209-3217.

Jevons, M. P. (1961). Celbenin-resistant staphylococci. *British Medical Journal, 1*, 124-125.

Joerger, A. C., Mayer, S. & Fersht, A. R. (2003). Mimicking natural evolution *in vitro*: an *N*-acetylneuraminate lyase mutant with an increased dihydrodipicolinate synthase activity. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 5694-5699.

Johnston, J. W., Zaleski, A., Allen, S., Mootz, J. M., Armbruster, D., Gibson, B. W., Apicella, M. A. & Munson, R. S., Jr. (2007). Regulation of sialic acid transport and catabolism in *Haemophilus influenzae*. *Molecular Microbiology, 66*, 26-39.

Jones, P. M. & George, A. M. (2004). The ABC transporter structure and mechanism: perspectives on recent research. *Cellular and Molecualr Life Science, 61*, 682-699.

Kalivoda, K. A., Steenbergen, S. M. & Vimr, E. R. (2013). Control of the *Escherichia coli* sialoregulon by transcriptional repressor NanR. *Journal of Bacteriology, 195*, 4689-4701.

Kalivoda, K. A., Steenbergen, S. M., Vimr, E. R. & Plumbridge, J. (2003). Regulation of sialic acid catabolism by the DNA binding protein NanR in *Escherichia coli*. *Journal of Bacteriology, 185*, 4806-4815.

Kelly, D. J. & Thomas, G. H. (2001). The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *Federation of European Microbiological Societies Microbiology Reviews, 25*, 405-424.

Kim, B. S., Hwang, J., Kim, M. H. & Choi, S. H. (2011). Cooperative regulation of the *Vibrio vulnificus* nan gene cluster by NanR protein, cAMP receptor protein, and *N*-acetylmannosamine 6-phosphate. *Journal of Biological Chemistry, 286*, 40889-40899.

Larion, M., Moore, L. B., Thompson, S. M. & Miller, B. G. (2007). Divergent evolution of function in the ROK sugar kinase superfamily: role of enzyme loops in substrate specificity. *Biochemistry, 46*, 13564-13572.

Lawrence, M., Barbosa, J., Smith, B., Hall, N., Pilling, P., Ooi, H. & Marcuccio, S. (1997). Structure and mechanism of a sub-family of enzymes related to *N*-acetylneuraminate lyase. *Journal of Molecular Biology, 266*, 381-399.

Liu, C., Li, D., Liang, Y. H., Li, L. F. & Su, X. D. (2008). Ring-opening mechanism revealed by crystal structures of NagB and its ES intermediate complex. *Journal of Molecular Biology, 379*, 73-81.

Locher, K. P., Lee, A. T. & Rees, D. C. (2002). The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science, 296*, 1091-1098.

Lowy, F. D. (2003). Antimicrobial resistance: the example of *Staphylococcus aureus*. *The Journal of Clinical Investigation, 111*, 1265-1273.

Marion, C., Aten, A. E., Woodiga, S. A. & King, S. J. (2011a). Identification of an ATPase, MsmK, which energises multiple carbohydrate ABC transporters in *Streptococcus pneumoniae*. *Infection and Immunity, 79*, 4193-4200.

Marion, C., Burnaugh, A. M., Woodiga, S. A. & King, S. J. (2011b). Sialic acid transport contributes to pneumococcal colonisation. *Infection and Immunity, 79*, 1262-1269.

Martinez, J., Nguyen, L. D., Hinderlich, S., Zimmer, R., Tauberger, E., Reutter, W., Saenger, W., Fan, H. & Moniot, S. (2012). Crystal structures of *N*-acetylmannosamine kinase provide insights into enzyme activity and inhibition. *The Journal of Biological Chemistry, 287*, 13656-13665.

Martinez, J., Steenbergen, S. & Vimr, E. (1995). Derived structure of the putative sialic acid transporter from *Escherichia coli* predicts a novel sugar permease domain. *Journal of Bacteriology, 177*, 6005-6010.

McDonald, M., Hurse, A. & Sim, K. N. (1981). Methicillin-resistant *Staphylococcus aureus* bacteraemia. *Medical Journal of Australia, 2*, 191-194.

Mirwaldt, C., Korndorfer, I. & Huber, R. (1995). The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. *Journal of Molecular Biology, 246*, 227-239.

Mulligan, C., Fischer, M. & Thomas, G. H. (2011). Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *Federation of European Microbiological Societies Microbiology Reviews, 35*, 68-86.

Mulligan, C., Geertsma, E. R., Severi, E., Kelly, D. J., Poolman, B. & Thomas, G. H. (2009). The substrate-binding protein imposes directionality on an electrochemical sodium gradient-driven TRAP transporter. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 1778-1783.

Mulligan, C., Leech, A. P., Kelly, D. J. & Thomas, G. H. (2012). The membrane proteins SiaQ and SiaM form an essential stoichiometric complex in the sialic acid tripartite ATP-independent periplasmic (TRAP) transporter SiaPQM (VC1777–1779) from *Vibrio cholerae*. *Journal of Biological Chemistry, 287*, 3598-3608.

North, R. A., Kessans, S. A., Griffin, M. D., Watson, A. J., Fairbanks, A. J. & Dobson, R. C. (2014a). Cloning, expression, purification, crystallisation and preliminary X-ray diffraction analysis of *N*-acetylmannosamine-6-phosphate 2-epimerase from methicillin-resistant *Staphylococcus aureus*. *Acta crystallographica Section F, Structural Biology and Crystallisation Communications, 70*, 650-655.

North, R. A., Seizova, S., Stampfli, A., Kessans, S. A., Suzuki, H., Griffin, M. D., Kvansakul, M. & Dobson, R. C. (2014b). Cloning, expression, purification, crystallisation and preliminary X-ray diffraction analysis of *N*-acetylmannosamine kinase from methicillin-resistant *Staphylococcus aureus*. *Acta crystallographica Section F, Structural Biology and Crystallisation Communications, 70*, 643-649.

Oldham, M. L., Khare, D., Quiocho, F. A., Davidson, A. L. & Chen, J. (2007). Crystal structure of a catalytic intermediate of the maltose transporter. *Nature, 450*, 515-521.

Oliva, G., Fontes, M. R., Garratt, R. C., Altamirano, M. M., Calcagno, M. L. & Horjales, E. (1995). Structure and catalytic mechanism of glucosamine 6-phosphate deaminase from *Escherichia coli* at 2.1 Å resolution. *Structure, 3*, 1323-1332.

Olson, M. E., King, J. M., Yahr, T. L. & Horswill, A. R. (2013). Sialic acid catabolism in *Staphylococcus aureus*. *Journal of Bacteriology, 195*, 1779-1788.

Panlilio, A. L., Culver, D. H., Gaynes, R. P., Banerjee, S., Henderson, T. S., Tolson, J. S. & Martone, W. J. (1992). Methicillin-resistant *Staphylococcus aureus* in U.S. hospitals, 1975-1991. *Infection Control and Hospital Epidemiology, 13*, 582-586.

Pelissier, M. C., Sebban-Kreuzer, C., Guerlesquin, F., Brannigan, J. A., Bourne, Y. & Vincent, F. (2014). Structural and functional characterisation of the *Clostridium perfringens N*-acetylmannosamine-6-phosphate 2-epimerase essential for the sialic acid salvage pathway. *Journal of Biological Chemistry, 289*, 35215-35224.

Poolman, B. & Konings, W. N. (1993). Secondary solute transport in bacteria. *Biochimica et Biophysica Acta, 1183*, 5-39.

Post, D. M., Mungur, R., Gibson, B. W. & Munson, R. S., Jr. (2005). Identification of a novel sialic acid transporter in *Haemophilus ducreyi*. *Infection and Immunity, 73*, 6727-6735.

Previato, J. O., Andrade, A. F., Pessolani, M. C. & Mendonca-Previato, L. (1985). Incorporation of sialic acid into *Trypanosoma cruzi* macromolecules. A proposal for a new metabolic route. *Molecular Biochemistry and Parasitology, 16*, 85-96.

Rudino-Pinera, E., Morales-Arrieta, S., Rojas-Trejo, S. P. & Horjales, E. (2002). Structural flexibility, an essential component of the allosteric activation in *Escherichia coli* glucosamine-6-phosphate deaminase. *Acta Crystallographica Section D, Biological Crystallography, 58*, 10-20.

Sauer, R. T., Yocum, R. R., Doolittle, R. F., Lewis, M. & Pabo, C. O. (1982). Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. *Nature, 298*, 447-451.

Saurin, W., Hofnung, M. & Dassa, E. (1999). Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *Journal of Molecular Evolution, 48*, 22-41.

Schauer, R. (2000). Achievements and challenges of sialic acid research. *Glycoconjugate Journal, 17*, 485-499.

Schauer, R. (2004). Sialic acids: fascinating sugars in higher animals and man. *Zoology, 107*, 49-64.

Severi, E., Hood, D. W. & Thomas, G. H. (2007). Sialic acid utilisation by bacterial pathogens. *Microbiology, 153*, 2817-2822.

Severi, E., Hosie, A. H. F., Hawkhead, J. A. & Thomas, G. H. (2010). Characterisation of a novel sialic acid transporter of the sodium solute symporter (SSS) family and *in vivo* comparison with known bacterial sialic acid transporters. *Federation of European Microbiological Societies Microbiology Letters, 304*, 47-54.

Severi, E., Randle, G., Kivlin, P., Whitfield, K., Young, R., Moxon, R., Kelly, D., Hood, D. & Thomas, G. H. (2005). Sialic acid transport in *Haemophilus influenzae* is essential for lipopolysaccharide sialylation and serum resistance and is dependent on a novel tripartite ATP-independent periplasmic transporter. *Molecular Microbiology, 58*, 1173-1185.

Shakhnovich, E. A., King, S. J. & Weiser, J. N. (2002). Neuraminidase expressed by *Streptococcus pneumoniae* desialylates the lipopolysaccharide of N*eisseria meningitidis* and *Haemophilus influenzae*: a paradigm for interbacterial competition among pathogens of the human respiratory tract. *Infection and Immunity, 70*, 7161-7164.

Shell, D. M., Chiles, L., Judd, R. C., Seal, S. & Rest, R. F. (2002). The neisseria lipooligosaccharide-specific alpha-2,3-sialyltransferase is a surface-exposed outer membrane protein. *Infection and Immunity, 70*, 3744-3751.

Shintre, C. A., Pike, A. C., Li, Q., Kim, J. I., Barr, A. J., Goubin, S., Shrestha, L., Yang, J., Berridge, G., Ross, J., Stansfeld, P. J., Sansom, M. S., Edwards, A. M., Bountra, C., Marsden, B. D., von Delft, F., Bullock, A. N., Gileadi, O., Burgess-Brown, N. A. & Carpenter, E. P. (2013). Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 9710-9715.

Sillanaukee, P., Ponnio, M. & Jaaskelainen, I. P. (1999). Occurrence of sialic acids in healthy humans and different disorders. *European Journal of Clinical Investigation, 29*, 413-425.

Sohanpal, B. K., El-Labany, S., Lahooti, M., Plumbridge, J. A. & Blomfield, I. C. (2004). Integrated regulatory responses of fimB to *N*-acetylneuraminic (sialic) acid and

GlcNAc in *Escherichia coli* K-12. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 16322-16327.

Sohanpal, B. K., Friar, S., Roobol, J., Plumbridge, J. A. & Blomfield, I. C. (2007). Multiple co-regulatory elements and IHF are necessary for the control of fimB expression in response to sialic acid and *N*-acetylglucosamine in *Escherichia coli* K-12. *Molecular Microbiology, 63*, 1223-1236.

Steenbergen, S. M., Lichtensteiger, C. A., Caughlan, R., Garfinkle, J., Fuller, T. E. & Vimr, E. R. (2005). Sialic acid metabolism and systemic pasteurellosis. *Infection and Immunity, 73*, 1284-1294.

Tanner, M. E. (2005). The enzymes of sialic acid biosynthesis. *Bioorganic Chemistry, 33*, 216-228.

Timms, N., Windle, C. L., Polyakova, A., Ault, J. R., Trinh, C. H., Pearson, A. R., Nelson, A. & Berry, A. (2013). Structural insights into the recovery of aldolase activity in *N*-acetylneuraminic acid lyase by replacement of the catalytically active lysine with gamma-thialysine by using a chemical mutagenesis strategy. *Chembiochem: a European Journal of Chemical Biology, 14*, 474-481.

Toner, E., Adalja, A., Gronvall, G. K., Cicero, A. & Inglesby, T. V. (2015). Antimicrobial resistance is a global health emergency. *Health Security, 13*, 153-155.

Varki, A. (1992). Diversity in the sialic acids. *Glycobiology, 2*, 25-40.

Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology, 3*, 97-130.

Vimr, E. & Lichtensteiger, C. (2002). To sialylate, or not to sialylate: that is the question. *Trends in Microbiology, 10*, 254-257.

Vimr, E. R. (1994). Microbial sialidases: does bigger always mean better? *Trends in Microbiology, 2*, 271-277.

Vimr, E. R. (2013). Unified theory of bacterial sialometabolism: how and why bacteria metabolise host sialic acids. *International Scholarly Research Notices Microbiology, 2013*, 816713.

Vimr, E. R., Kalivoda, K. A., Deszo, E. L. & Steenbergen, S. M. (2004). Diversity of microbial sialic acid metabolism. *Microbiology and Molecular Biology Reviews, 68*, 132-153.

Vimr, E. R. & Troy, F. A. (1985). Identification of an inducible catabolic system for sialic acids (*nan*) in *Escherichia coli*. *Journal of Bacteriology, 164*, 845-853.

Vincent, F., Davies, G. J. & Brannigan, J. A. (2005). Structure and kinetics of a monomeric glucosamine 6-phosphate deaminase: missing link of the NagB superfamily? *Journal of Biological Chemistry, 280*, 19649-19655.

Vincent, F., Yates, D., Garman, E., Davies, G. J. & Brannigan, J. A. (2004). The three-dimensional structure of the *N*-acetylglucosamine-6-phosphate deacetylase, NagA, from *Bacillus subtilis*: a member of the urease superfamily. *Journal of Biological Chemistry, 279*, 2809-2816.

Vogel-Scheel, J., Alpert, C., Engst, W., Loh, G. & Blaut, M. (2010). Requirement of purine and pyrimidine synthesis for colonisation of the mouse intestine by *Escherichia coli*. *Applied and Environmental Microbiology, 76*, 5181-5187.

von Itzstein, M. (2007). The war against influenza: discovery and development of sialidase inhibitors. *Nature Reviews Drug Discovery, 6*, 967-974.

Wirth, C., Condemine, G., Boiteux, C., Berneche, S., Schirmer, T. & Peneff, C. M. (2009). NanC crystal structure, a model for outer-membrane channels of the acidic sugar-specific KdgM porin family. *Journal of Molecular Biology, 394*, 718-731.

Woo, J. S., Zeltina, A., Goetz, B. A. & Locher, K. P. (2012). X-ray structure of the *Yersinia pestis* heme transporter HmuUV. *Nature Structural and Molecular Biology, 19*, 1310-1315.

Wright, E. M., Loo, D. D. F., Hirayama, B. A. & Turk, E. (2004). Surprising versatility of Na^+-glucose cotransporters: SLC5. *Physiology, 19*, 370-376.

Chapter Two

The structure, function and inhibition of methicillin-resistant *Staphylococcus aureus N*-acetylneuraminate lyase

Part of the work presented in this chapter has been published in:

North, R. A., Kessans, S. A., Atkinson, S. A., Suzuki, H., Watson, A. J. A., Burgess, B., Angley, L., Hudson, A. O., Varsani, A., Griffin, M. D. W., Fairbanks, A. J., Dobson, R. C. J. (2013). Cloning, expression, purification, crystallisation and preliminary X-ray diffraction studies of *N*-acetylneuraminate lyase from methicillin resistant *Staphylococcus aureus*. *Acta Crystallographica Section F, Structural Biology and Crystallisation Communications*, *69*, 306-312.

2.1 Introduction

2.1.1 *N*-Acetylneuraminate lyase is a target for drug design

In low glucose and heavily sialylated niches, such as the human respiratory and gastrointestinal tract, highly abundant sialic acid molecules provide an alternate source of carbon, nitrogen and energy for bacterial pathogens residing within the host (Martinez *et al.*, 1995; Chang *et al.*, 2004; Vimr *et al.*, 2004; Severi *et al.*, 2005). Some bacterial pathogens have evolved a mechanism by which they can degrade these host-derived sialic acid into fructose-6-phosphate, a metabolite necessary for the survival of such pathogens (Vimr *et al.*, 2004; Almagro-Moreno & Boyd, 2009a). As described in Chapter One,

various *in vitro* and *in vivo* studies have implicated the sialic acid degradation pathway as a novel candidate for antibiotic drug development (Chang *et al.*, 2004; Almagro-Moreno & Boyd, 2009b; Jeong *et al.*, 2009; Pezzicoli *et al.*, 2012; Olson *et al.*, 2013). Following the import of sialic acid from host to pathogen, the first and committed enzyme involved in sialic acid degradation is known as *N*-acetylneuraminate lyase. This enzyme has been recognised as a viable antibiotic drug target (Severi *et al.*, 2007; von Itzstein, 2007). Deletion of the *nan*A gene coding for *N*-acetylneuraminate lyase in *Escherichia coli* (Vimr & Troy, 1985; Vogel-Scheel *et al.*, 2010), *Staphylococcus aureus* (Olson *et al.*, 2013), *Vibrio cholerae* (Almagro-Moreno & Boyd, 2009b) and *Vibrio vulnificus* (Jeong *et al.*, 2009) renders the organisms incapable of growing on sialic acid *in vitro*. Moreover, the *nan*A deletion strains of *V. cholerae* and *V. vulnificus* were defective for intestinal colonisation in mouse models, indicative of an impaired ability to colonise, persist and survive within the mouse intestine (Almagro-Moreno & Boyd, 2009b; Jeong *et al.*, 2009). *N*-Acetylneuraminate lyase therefore plays a significant role in the virulence inflicted by a range of human bacterial pathogens and understanding the structure and function of this enzyme will underpin the design of inhibitors that specifically target this enzyme.

2.1.2 The reaction catalysed by *N*-acetylneuraminate lyase

N-Acetylneuraminate lyase is an enzyme belonging to the aldolase class I superfamily, which shares a common structural framework to that of related triose phosphate isomerase barrel [$(\beta/\alpha)_8$-barrel] enzymes, including 4-hydroxy-tetrahydrodipicolinic acid synthase [formally known as dihydrodipicolinate synthase (DHDPS)] (Mirwaldt *et al.*, 1995; Dobson *et al.*, 2005; Burgess *et al.*, 2008; Kefala *et al.*, 2008), D-5-keto-4-deoxyglucarate dehydratase (Jeffcoat *et al.*, 1969b; Jeffcoat *et al.*, 1969a), 2-keto-3-deoxygluconate aldolase (Buchanan *et al.*, 1999), *trans*-*o*-hydroxybenzylidenepyruvate hydratase-aldolase (Eaton, 1994) and *trans*-2'-carboxybenzalpyruvate hydratase-aldolase (Iwabuchi & Harayama, 1998). Although these enzymes all share a common structural framework, they are responsible for catalysing different reactions on separate biochemical pathways. *N*-Acetylneuraminate lyase catalyses the retro-aldol cleavage of open chain sialic acid, which yields *N*-acetylmannosamine and pyruvate (Figure 2.1) (Izard *et al.*, 1994). The equilibrium of this reaction strongly favours the cleavage of sialic acid, although the reverse aldol condensation of *N*-acetylmannosamine and pyruvate to form sialic acid is a

convenient route for the synthesis of sialic acids (Lilley *et al.*, 1992; Li *et al.*, 2008; Campeotto *et al.*, 2009; Campeotto *et al.*, 2010; Huynh *et al.*, 2013).

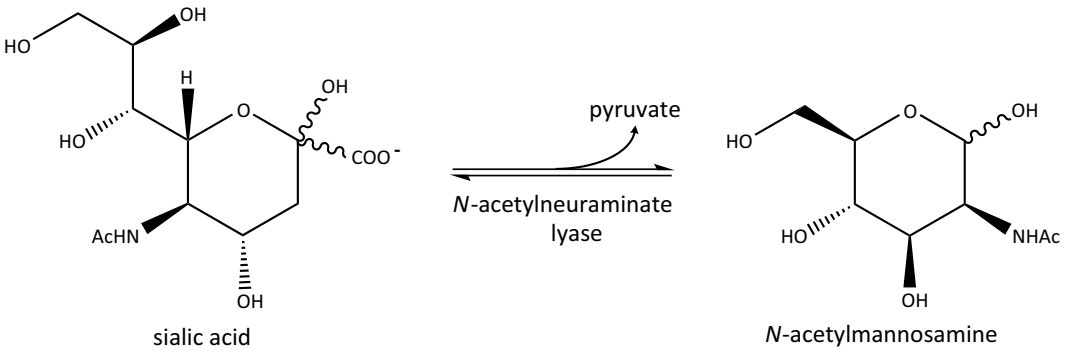


Figure 2.1. The reaction catalysed by *N*-acetylneuraminate lyase. Sialic acid is cleaved to yield *N*-acetylmannosamine and pyruvate.

2.1.3 The catalytic mechanism of *N*-acetylneuraminate lyase

Although members of the *N*-acetylneuraminate lyase sub-family of enzymes catalyse unique reactions on separate biochemical pathways, these enzymes share a unifying step in their reaction pathway, which involves the formation of a Schiff base between the amine nitrogen of a strictly conserved lysine residue and the C2 carbon of an α-keto acid moiety on respective substrates (Izard *et al.*, 1994). The first X-ray structural analyses of *E. coli N*-acetylneuraminate lyase (Lawrence *et al.*, 1997) and *E. coli* DHDPS (Blickling *et al.*, 1997) in complex with pyruvate and analogues thereof supported this notion.

These structures allowed for residues involved in the binding of pyruvate and the associated aldol cleavage/condensation step to be identified. For the *E. coli N*-acetylneuraminate lyase in particular, a highly conserved Lys165 residue is responsible for the formation of a Schiff base with the C2 carbon of the α-keto acid moiety of open chain sialic acid in the retro-aldol direction of catalysis. Hydrogen bonds are also formed between the carboxylate moiety of the substrate and the hydroxyl of Tyr137, as well as the backbone amides of Ser47 and Thr48, which form the respective second and third residues of a highly conserved GxxGE motif (Lawrence *et al.*, 1997). A catalytic triad is formed between the side chains of Tyr137, Ser47 and Tyr110 from a neighbouring subunit. In

DHDPS enzymes, a similar catalytic triad acts as a proton shuttle during catalysis (Dobson *et al.*, 2004; Dobson *et al.*, 2005). Thus, the presence of this triad in *N*-acetylneuraminate lyase enzymes provides strong evidence that it also plays a key role in catalysis (Daniels *et al.*, 2014).

Mutagenesis, kinetic analysis and further structures of *N*-acetylneuraminate lyase enzymes in complex with sialic acid or sialic acid analogues have provided additional insight into residues involved in substrate binding and the chemistry that is catalysed (Figure 2.2) (Barbosa *et al.*, 2000; Kruger *et al.*, 2001; Huynh *et al.*, 2013; Daniels *et al.*, 2014). In *E. coli N*-acetylneuraminate lyase in particular, residues Thr167, Gly189, Asp191, Glu192 and Ser208 have also been implicated in substrate binding (Daniels *et al.*, 2014). Importantly, all residues demonstrated to be involved in catalysis are highly conserved between bacterial species within the *N*-acetylneuraminate lyase sub-family of enzymes, illustrating the importance of these residues to the structure and function of these enzymes (Blickling *et al.*, 1997; Lawrence *et al.*, 1997; Barbosa *et al.*, 2000; Kruger *et al.*, 2001; Huynh *et al.*, 2013; North *et al.*, 2013).

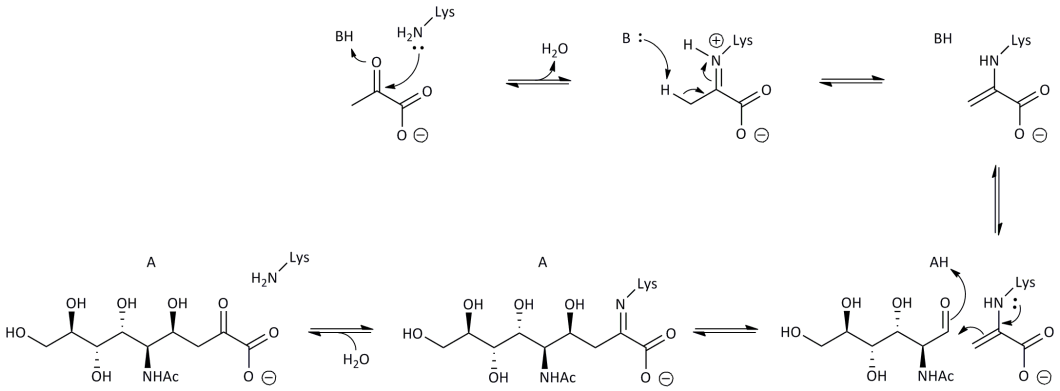


Figure 2.2. The proposed catalytic mechanism of *N*-acetylneuraminate lyase. The enzymatic reaction showing the involvement of a conserved lysine residue in catalysis. The lysine residue is involved in the formation of a Schiff base with the C2 carbon of the α-keto acid moiety of open chain sialic acid in the retro-aldol sense and pyruvate in the aldol direction.

2.1.4 The structure of *N*-acetylneuraminate lyase

Various structures of *N*-acetylneuraminate lyase enzymes have been solved and studied, including *N*-acetylneuraminate lyase from *E. coli* (Izard *et al.*, 1994; Lawrence *et al.*, 1997; Joerger *et al.*, 2003; Campeotto *et al.*, 2009; Campeotto *et al.*, 2010; Daniels *et al.*, 2014), *Haemophilus influenzae* (Barbosa *et al.*, 2000), *Pasteurella multocida* (Huynh *et al.*, 2013) and *S. aureus* (Timms *et al.*, 2013). In all structures published, *N*-acetylneuraminate lyase has been shown to be a tetramer with four identical $(\beta/\alpha)_8$-barrel monomers, each followed by an extension of three α-helices at the C-terminus (Izard *et al.*, 1994). Similar to DHDPS, the tetrameric structure of *E. coli N*-acetylneuraminate lyase can be described as a dimer of dimers, where monomers *a* and *b* form one dimer and monomers *c* and *d* form the other (Figure 2.3 A). The association between monomers leaves a large pore in the central cavity of the tetramer.

Within the monomer of *E. coli N*-acetylneuraminate lyase, the consecutive β-strands are labelled *a* through *h* and the α-helices are labelled *A* through *K* with the three additional C-terminal α-helices denoted as *I*, *J*, and *K*, respectively (Figure 2.3 B) (Izard *et al.*, 1994). As with all $(\beta/\alpha)_8$-barrel enzymes, the active site is located in the centre of the barrel at the C-terminal end of the β-strands (Wierenga, 2001). For *N*-acetylneuraminate lyase, these C-terminal ends of the β-strands are oriented inward, toward the central cavity of the tetramer.

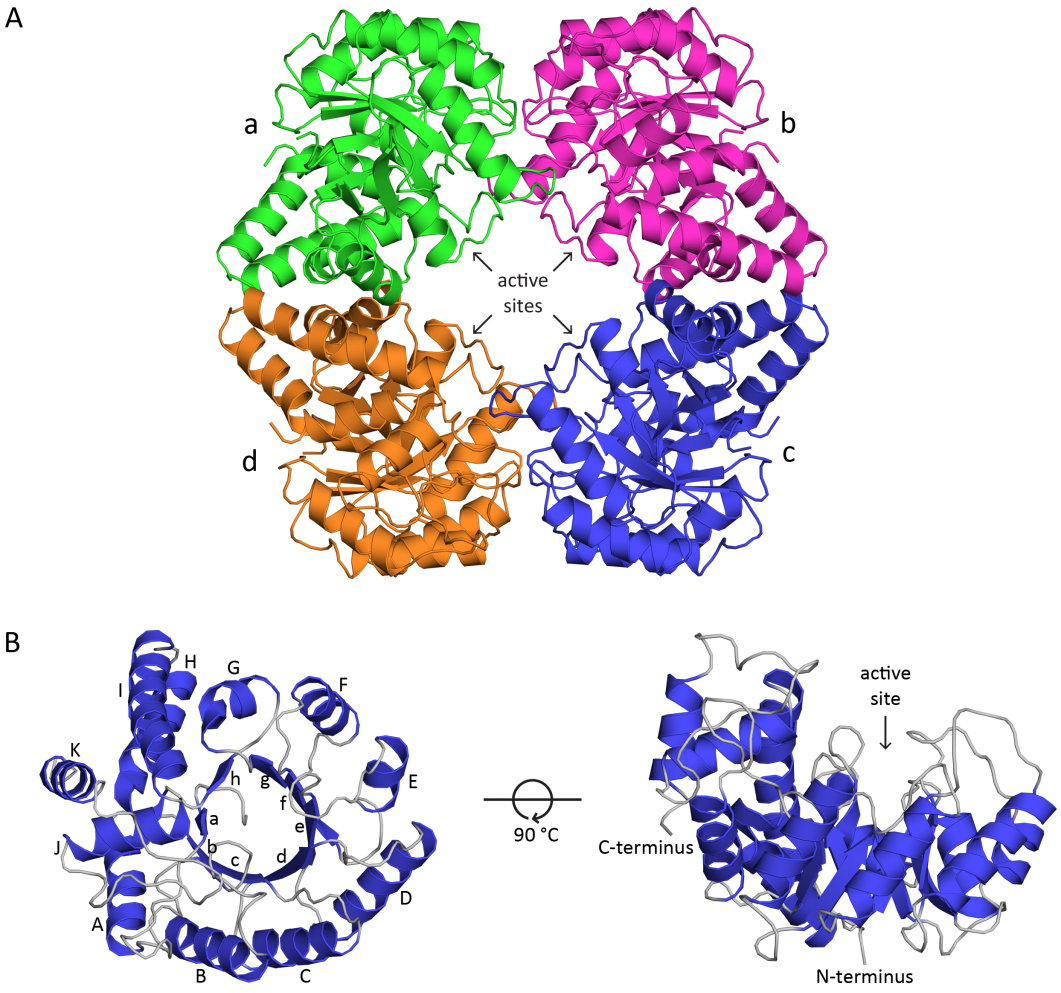


Figure 2.3. The structure of *E. coli N*-acetylneuraminate lyase. (A) The tetrameric structure of *E. coli N*-acetylneuraminate lyase. Monomers are labelled *a* through *d*. The active site of each monomer is oriented towards the central cavity of the tetramer. **(B)** The monomeric structure of *E. coli N*-acetylneuraminate lyase. The β-strands are labelled *a* through *h* and the α-helices are labelled *A* through *K*. The active site is located in the centre of the barrel at the C-terminal end of the β-strands.

2.1.5 Overview of this chapter

This chapter will address the structure, function and inhibition of *N*-acetylneuraminate lyase from methicillin-resistant *S. aureus* (MRSA), a pathogen that causes significant disease globally (Grundmann *et al.*, 2006; Chambers & Deleo, 2009). The overarching aim of this chapter is to investigate the structure and function of MRSA *N*-acetylneuraminate lyase with a first generation inhibitor that targets this enzyme. This chapter presents a

solution based characterisation of MRSA *N*-acetylneuraminate lyase to ensure that the enzyme was stable and folded. In addition, the quaternary structure was investigated, which is important for the interaction study presented in Chapter Four. A kinetic and structural analysis of MRSA *N*-acetylneuraminate lyase with and without a successful first generation inhibitor is presented and the mechanism of inhibition is analysed. Although the structure of *S. aureus N*-acetylneuraminate lyase (PDB entry 4ahp) (Timms *et al.*, 2013) was solved and published during the course of this study, the structure of MRSA *N*-acetylneuraminate lyase presented in this thesis is of a significantly higher resolution and contains information that is missing from PDB entry 4ahp. Combined, this type of data is important for further optimisation of inhibitors, leading to the development of novel antibiotic drugs that target the clinically important human pathogen, MRSA.

2.2 Results and discussion

2.2.1 Purification of MRSA *N*-acetylneuraminate lyase

The *nanA* gene encoding *N*-acetylneuraminate lyase from MRSA252 (Appendix) was supplied in a pET11a expression plasmid. The construct was transformed into *E. coli* BL21(DE3) bacterial cells, cultured and expressed in Luria Bertani medium, harvested, and lysed (Chapter Six, Sections 6.6.1-6.6.3). This construct was subjected to purification by anion exchange chromatography, hydrophobic interaction chromatography and size exclusion chromatography (Chapter Six, Section 6.6.8). The increasing purity of MRSA *N*-acetylneuraminate lyase from each purification procedure is shown in Figure 2.4. The resulting purity was visually estimated to be at least 95%, which can be observed by the single band at 33 kilodalton (kDa) following size exclusion chromatography (Figure 2.4, lane 4). The mass of MRSA *N*-acetylneuraminate lyase was confirmed by electrospray ionisation mass spectrometry following purification.

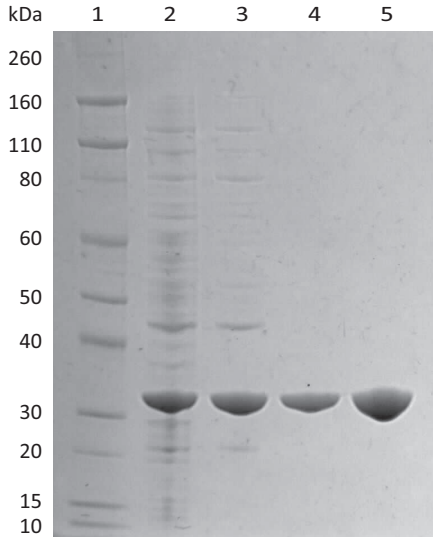


Figure 2.4. Purification of MRSA *N*-acetylneuraminate lyase. Sodium dodecyl sulfate polyacrylamide gel electrophpresis (SDS-PAGE) analysis showing the increasing purity of MRSA *N*-acetylneuraminate lyase at each purification step. Lane 1, protein ladder (kDa); lane 2, crude; lane 3, pooled fractions from anion exchange chromatography; lane 4, pooled fractions from hydrophobic interaction chromatography; lane 5, post size exclusion chromatography.

2.2.2 Solution structure of MRSA *N*-acetylneuraminate lyase

2.2.2.1 Secondary structure as determined by circular dichroism spectroscopy

To ensure that MRSA *N*-acetylneuraminate lyase was folded, the secondary structure was examined by circular dichroism (CD) spectroscopy. The spectrum recorded showed that the enzyme is folded (Figure 2.5). A double minimum is apparent at 208 and 222 nm. This was also observed for the closely related DHDPS enzymes (Burgess *et al.*, 2008; Griffin *et al.*, 2008).

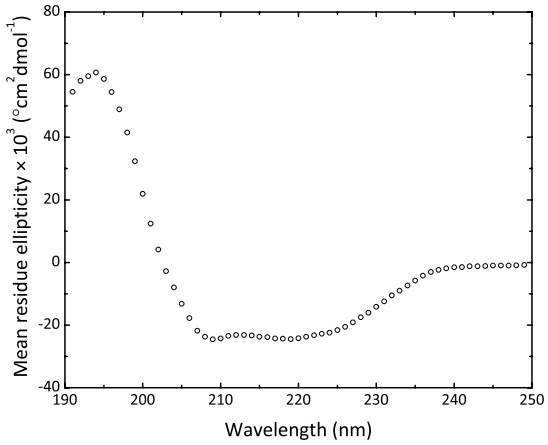


Figure 2.5. CD spectrum of MRSA *N*-acetylneuraminate lyase. CD spectroscopy was conducted at 0.1 mg/mL in 20 mM 2-amino-2-hydroxymethyl-propane-1,3-diol (Tris), pH 8.0. Mean residue ellipticity is plotted as a function of wavelength. Data are represented as open circles (o).

2.2.2.2 Quaternary structure as determined by analytical ultracentrifugation

Sedimentation velocity experiments using analytical ultracentrifugation were used to characterise the quaternary structure of MRSA *N*-acetylneuraminate lyase in solution for the first time. Interestingly, the related MRSA DHDPS enzyme exists in a monomer-dimer equilibrium in solution (Burgess *et al.*, 2008), while nearly all other studied DHDPS and *N*-acetylneuraminate lyase enzymes share a tetrameric arrangement. Data were fitted to a continuous sedimentation coefficient distribution [c(*s*)] model, resulting in a distinct single, narrow and symmetrical peak with a sedimentation coefficient of 6.6 S (Figure 2.6 A). This confirms that MRSA *N*-acetylneuraminate lyase is a single species in solution.

This data was then fitted to a continuous mass distribution [c(*M*)] to calculate the apparent molecular mass of the species (Figure 2.6 B), which was determined to be 133 kDa. Considering the theoretical monomeric mass of MRSA *N*-acetylneuraminate lyase is 33 kDa, this is consistent with a theoretical tetrameric mass of 129 kDa. The absorbance at 280 nm versus radial position of MRSA *N*-acetylneuraminate lyase shows a single sedimenting boundary, consistent with the presence of a single species and the randomly distributed residuals indicate a good fit to the data (Figure 2.6 C) (Lebowitz *et al.*, 2002).

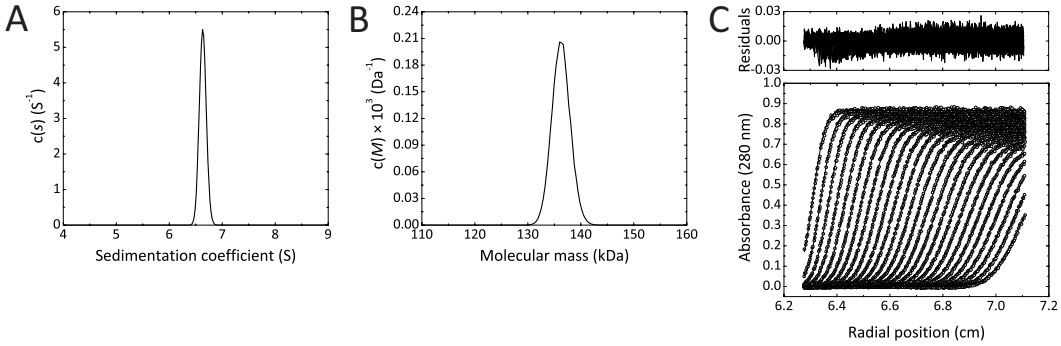


Figure 2.6. Sedimentation velocity analysis of MRSA *N*-acetylneuraminate lyase using analytical ultracentrifugation. The experiment was conducted at 0.5 mg/mL (15.1 µM) of freshly purified protein in 20 mM Tris, pH 8.0, 150 mM sodium chloride. Data were collected at 280 nm, 50 000 rpm and 20 °C. *SEDNTERP* was used to calculate the partial specific volume of MRSA *N*-acetylneuraminate lyase (0.7404 g/ml), solvent density (1.0029 g/ml) and viscosity (0.0102 poise) (Laue *et al.*, 1992). Data were fitted to a c(*s*) and c(*M*) model at a resolution of 300 and a confidence level of 0.95 using *SEDFIT* (Schuck, 2000). **(A)** The c(*s*) model plotted as a function of sedimentation coefficient. Data were fitted with an *s* value ranging between 4 S and 10 S. The fit resulted in a frictional ratio (f/f_0) of 1.3394. **(B)** The c(*M*) model plotted as a function of molecular mass (kDa). Data were fitted with a minimum mass of 20 kDa and a maximum mass of 180 kDa. **(C)** Absorbance at 280 nm plotted as a function of radial position (cm). The raw data are represented as open symbols (o) and overlaid with the non-linear least squares best fit. The residuals for this fit are shown above.

2.2.2.3 Small angle X-ray scattering analysis of the solution structure

Small angle X-ray scattering (SAXS) data further verifies the solution structure of MRSA *N*-acetylneuraminate lyase determined by sedimentation velocity analysis. Sample and data quality were assessed by Guinier analysis of the SAXS scattering curve at low angles. The linearity of the Guinier plot (Figure 2.7 A) confirms that the scattering of MRSA *N*-

acetylneuraminate lyase is reliable and no sample aggregation and/or interparticle interference is present (Putnam *et al.*, 2007). Scattering data are presented as an intensity plot (Figure 2.7 B) and a real space distance distribution function *P*(*r*) plot (Figure 2.7 C). The maximum dimension of the scattering particle (D_{max}) was calculated by *AUTOPOROD* (Petoukhov & Svergun, 2007) as 88.7 Å. The radius of gyration (R_g) determined by *GNOM* (Svergun, 1992) was calculated as 32.8 Å, which is in agreement with the R_g determined by Guinier analysis. The molecular mass for MRSA *N*-acetylneuraminate lyase extracted from the SAXS data is 112 kDa, which is slightly smaller than the theoretical tetrameric mass of 129 kDa and the apparent molecular mass of 133 kDa derived from sedimentation velocity analysis. However, the resolution of SAXS data is low, which would account for this small discrepancy in size (Putnam *et al.*, 2007). The *P*(*r*) distribution shows a single peak that is symmetrical in shape, with no positively skewed tail at long distances. This indicates that MRSA *N*-acetylneuraminate lyase is globular in shape and not elongated in solution.

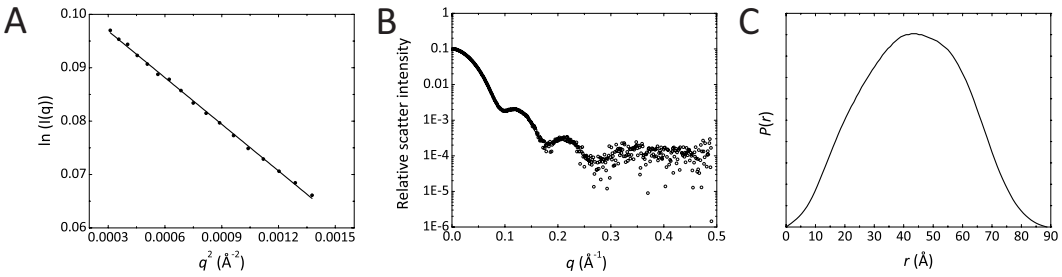


Figure 2.7. SAXS data for MRSA *N*-acetylneuraminate lyase. (A) A Guinier plot determined using *GNOM*, which is linear. **(B)** Experimental scattering data presented as an intensity plot. **(C) A** *P*(*r*) plot determined using *GNOM*.

It is important to directly compare SAXS curves with a crystal structure, since the crystal lattice can influence the observed structure (Putnam *et al.*, 2007). Thus, in an effort to determine whether the structure in solution accurately reflects my crystal structure (Section 2.2.7), the experimental scattering data were compared with the theoretical scattering generated from atomic coordinates of the crystal structure using *CRYSOL* (Svergun *et al.*, 1995). In addition, the experimental scattering was also compared with the theoretical scattering of the published *S. aureus N*-acetylneuraminate lyase structure (PDB entry 4ahp) (Timms *et al.*, 2013). The theoretical scattering calculated for my crystal structure of

MRSA *N*-acetylneuraminate lyase fits the experimental scattering data (χ^2 = 0.8), confirming that the crystal structure is an accurate representation of the solution based structure. Whereas, the theoretical scattering calculated for PDB entry 4ahp does not statistically fit the experimental scattering data (χ^2 = 1.9) (Figure 2.8). A poor fit is usually indicative of a difference between the protein in solution and the crystallised state (Putnam *et al.*, 2007). Even though these two structures share a 100% sequence identity, they have been crystallised in different conditions (Section 2.2.5). The difference may therefore be attributed to variable crystal packing between the structures (Section 2.2.9), or different protein dynamics in solution that are constrained in the crystal (Griffin *et al.*, 2008).

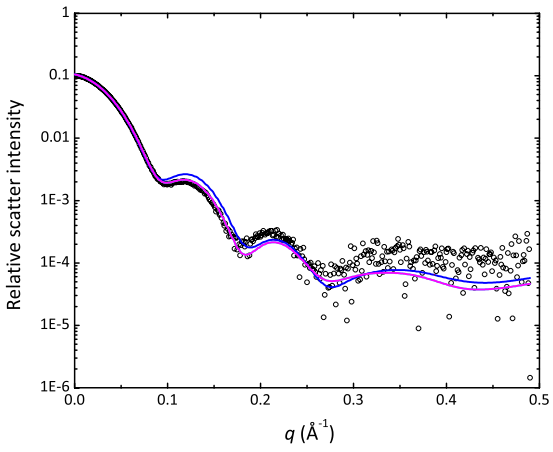


Figure 2.8. Comparison of the experimental scattering and theoretical scattering data. The experimental scattering profile of MRSA *N*-acetylneuraminate lyase (presented as open circles, o) overlaid with theoretical scattering profile of the MRSA *N*-acetylneuraminate lyase structure in magenta, and the theoretical scattering profile of the *S. aureus N*-acetylneuraminate lyase structure (PDB entry 4ahp) in blue.

2.2.3 Kinetic analysis of MRSA *N*-acetylneuraminate lyase

To characterise the kinetic properties of MRSA *N*-acetylneuraminate lyase, the absorbance of reduced nicotinamide adenine dinucleotide (NADH) at 340 nm was followed using a lactate dehydrogenase coupled assay (Chapter Six, Section 6.9.1) (Devenish & Gerrard, 2009). Briefly, the assay is initiated by the addition of substrate (sialic acid), which is converted into *N*-acetylmannosamine and pyruvate by MRSA *N*-acetylneuraminate lyase; lactate dehydrogenase (LDH) then converts the pyruvate into lactate whilst simultaneously

reducing NADH into the oxidised form of nicotinamide adenine dinucleotide (NAD+). Because NAD+ does not absorb at 340 nm, it is the reduction in absorbance at this wavelength that is followed spectrophotometrically, allowing the rate of catalysis by MRSA *N*-acetylneuraminate lyase to be followed.

2.2.3.1 Validity of a coupled enzyme assay

Prior to extensive kinetic analysis, various controls had to be performed to ensure that the conditions of the experiment allowed for accurate and reproducible measurements to be made. This means that although the assay is conducted under *in vitro* conditions, the results can be extrapolated to the expected activity occurring *in vivo* (Scopes, 2002). Firstly, in a coupled enzyme assay, it is important that the coupling enzyme is in excess so that it is not limiting the rate of the reaction being measured (Cornish-Bowden, 2004). This ensures that intermediate products do not accumulate and the coupled reactions can rapidly respond. For kinetic analysis of MRSA *N*-acetylneuraminate lyase, ensuring that the concentration of LDH is in excess meant that the rate of the reaction depended entirely upon MRSA *N*-acetylneuraminate lyase. The LDH concentration was tested between 3 and 400 μg per assay (Figure 2.9 A). Consistent with the concentration of LDH used previously (Devenish & Gerrard, 2009), a final amount of 100 μg per assay could be used as excess.

When utilising a coupled enzyme assay, it is also necessary that the rate of the reaction being measured is proportional to the amount of enzyme that is present (Cornish-Bowden, 2004). To test this, the initial velocity was analysed with 0.7 to 22.6 μg of MRSA *N*-acetylneuraminate lyase per assay. A final concentration of 1.4 μg was selected for the kinetic analysis of MRSA *N*-acetylneuraminate lyase, as this amount was proportional to the initial velocity, depicted by the linearity seen in Figure 2.9 B. In addition, this concentration resulted in an adequate amount of time for the reaction to reach completion, meaning that the initial velocity could be measured when less than 5% of the substrate had been depleted, minimising the effect of the reverse reaction, inhibition by a product or inactivation of the enzyme (Cornish-Bowden, 2004).

In addition, the effect of temperature on MRSA *N*-acetylneuraminate lyase activity was evaluated. This ensured that the temperatures selected to run the assays were appropriate. Previously published kinetic analyses of *N*-acetylneuraminate lyase enzymes were

conducted at 30 °C. However, inhibition studies have the potential to be translated into the development of antibiotic drugs, thus it is important that these analyses were performed at a physiological temperature of 37 °C. To test the effect of temperature on the function of MRSA *N*-acetylneuraminate lyase, the initial velocity of this enzyme between 20 °C and 80 °C was analysed (Figure 2.9 C). As depicted, MRSA *N*-acetylneuraminate lyase is stable up until approximately 70 °C, before the initial velocity begins to decline. Thus, it is safe to perform assays at a physiological temperature of 37 °C.

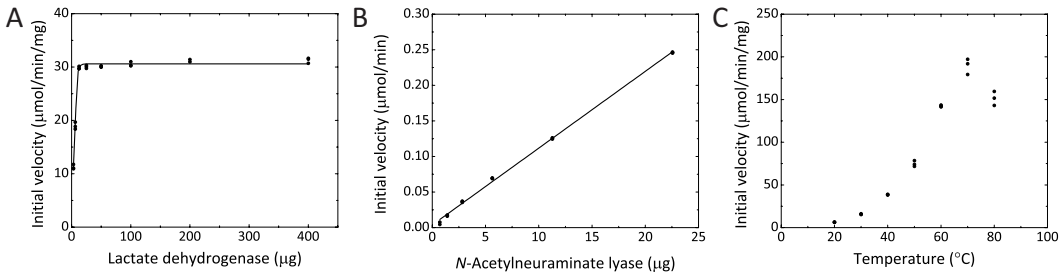


Figure 2.9. The initial velocity of MRSA *N*-acetylneuraminate lyase as a function of LDH concentration, the amount of enzyme present and temperature. (A) The initial velocity across a range of LDH concentrations. **(B)** The initial velocity in response to different concentrations of MRSA *N*-acetylneuraminate lyase. **(C)** The initial velocity plotted across a range of temperatures.

2.2.3.2 Michaelis-Menten kinetic analysis of MRSA N-*acetylneuraminate lyase*

The initial velocity of MRSA *N*-acetylneuraminate lyase with varying concentrations of sialic acid was measured and fitted to the Michaelis-Menten kinetic model, displaying a hyperbolic dependence of initial velocity on substrate concentration (Figure 2.10). The Michaelis-Menten constant (K_M) was calculated as 3.2 ± 0.1 mM, the maximum velocity (V_{max}) was calculated as 40.0 ± 0.2 µmol/min/mg and the catalytic turnover number (k_{cat}) was determined to be 22.1 ± 0.1 s^{-1}. These values are consistent with the kinetic parameters obtained for other *N*-acetylneuraminate lyase enzymes in the literature. For example, the *E. coli N*-acetylneuraminate lyase (Devenish & Gerrard, 2009) displayed a K_M of 4.1 mM and the *Clostridium perfringens N*-acetylneuraminate lyase (Kruger *et al.*, 2001) displayed a K_M of 3.2 mM. Further kinetic analyses of the *S. aureus* ((Timms *et al.*, 2013) and *E. coli* (Daniels *et al.*, 2014) *N*-acetylneuraminate lyase have been published by the same group,

however, the kinetic parameters presented are identical for both enzymes, which is unlikely. Thus, this may be the first kinetic analysis of MRSA *N*-acetylneuraminate lyase.

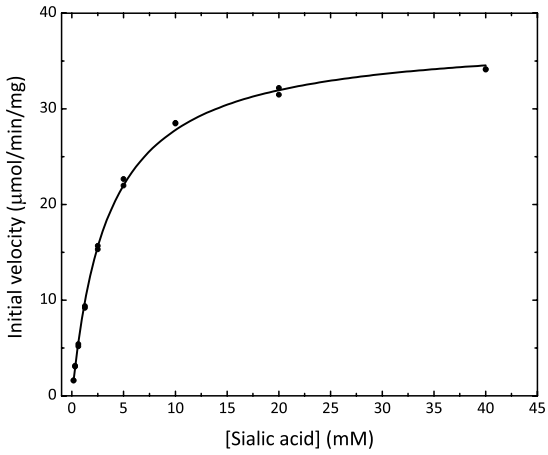


Figure 2.10. Michaelis-Menten kinetic analysis of MRSA *N*-acetylneuraminate lyase. Data were fitted to the Michaelis-Menten equation with an R^2 value of 0.99.

2.2.4 Inhibition of MRSA *N*-acetylneuraminate lyase

Analogues of sialic acid, including sialic acid alditol, 4-deoxy sialic acid and 4-oxo sialic acid have been shown to inhibit *N*-acetylneuraminate lyase to varying degrees. For comparison, the chemical structure of sialic acid is shown in Figure 2.11 A. The sialic acid alditols (Figure 2.11 B) have a hydroxyl group at the C2 position, as opposed to an aldehyde group, which is required for Schiff base formation with a conserved lysine residue (Deijl & Vliegenthart, 1983; Barbosa *et al.*, 2000; Ooi *et al.*, 2000). Conversely, 4-deoxy sialic acid (Figure 2.11 C) (Hagedorn & Brossmer, 1986; Zbiral *et al.*, 1992; Ooi *et al.*, 1999) and 4-oxo sialic acid (Figure 2.11 D) (Gross & Brossmer, 1988) are capable of Schiff base formation, but because they do not have a hydroxyl group at the C4 position, they cannot undergo aldol cleavage (Barbosa *et al.*, 2000). The inhibition constant (K_i) for the sialic acid alditols is 4.1 mM (Deijl & Vliegenthart, 1983), for 4-deoxy sialic acid it is 0.90 mM (Zbiral *et al.*, 1992) and for 4-oxo sialic acid it is 0.03 mM against *C. perfringens* *N*-acetylneuraminate lyase (Gross & Brossmer, 1988). The sialic acid alditols were synthesised in Professor Antony Fairbanks' laboratory (University of Canterbury) as a racemic mixture of the 2*R*- and 2*S*-diastereoisomers (Ooi *et al.*, 2000) and its ability to

inhibit MRSA *N*-acetylneuraminate lyase was tested by kinetic analysis. This type of data is important for further optimisation of inhibitors that target this bacterial pathogen.

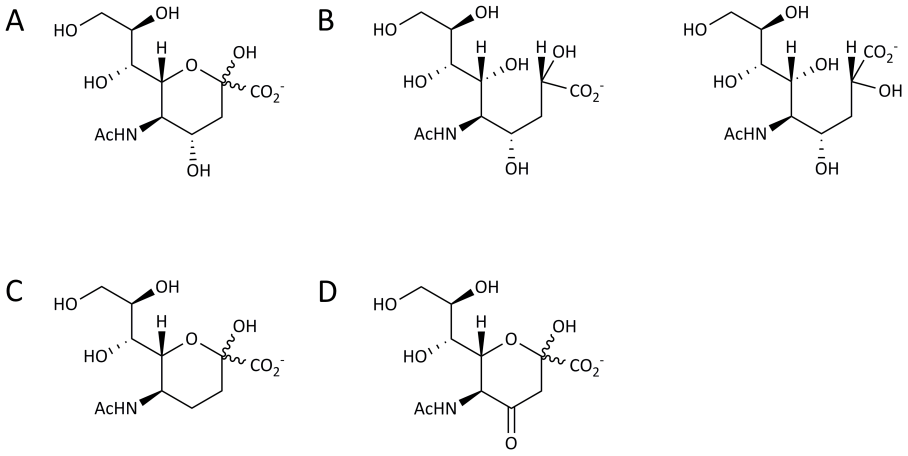


Figure 2.11. The chemical structure of *N*-acetylneuraminate lyase inhibitors. (A) The chemical structure of sialic acid for comparison. **(B)** The chemical structure of the 2*R*-diastereoisomer (left) and 2*S*-diastereoisomer (right) of sialic acid alditol. **(C)** The chemical structure of 4-deoxy sialic acid. **(D)** The chemical structure of 4-oxo sialic acid. Carbons are numbered one through to nine.

To determine the K_i for the sialic acid alditol mixture, the initial velocity of MRSA *N*-acetylneuraminate lyase with 0.16 to 20 mM of sialic acid against 0 to 5 mM of the sialic acid alditols was measured. These data were fitted to the competitive model of inhibition, non-competitive model of inhibition, un-competitive model of inhibition and the mixed model of inhibition. The kinetic parameters and R^2 value for these fits are shown in Table 2.1. The competitive inhibition model provided the best fit (Figure 2.12) and yielded a K_i of 0.39 mM ± 0.01. Interestingly, the K_i for the sialic acid alditols against MRSA *N*-acetylneuraminate lyase is considerably less than that demonstrated for the *C. perfringens* *N*-acetylneuraminate lyase (4.1 mM). Despite having identical active site residues in regards to those that are involved in catalysis (Kruger *et al.*, 2001) and the highly conserved active site configuration seen in *N*-acetylneuraminate lyase crystal structures, there appears to be species-specific inhibition of *N*-acetylneuraminate lyase enzymes from different bacterial pathogens. Consistent with Michaelis-Menten kinetic analysis, the K_M and V_{max} for the competitive inhibition model was determined to be 3.7 ± 0.1 mM and 37.8 ± 0.4 µmol/min/mg, respectively.

Table 2.1. Kinetic parameters for the inhibition of MRSA *N*-acetylneuraminate lyase. The K_i, K_M, V_{max} and R^2 for each fit were calculated in OriginPro.

Model of inhibition	K_i (mM)	K_M (mM)	V_{max} (µmol/min/mg)	R^2
Competitive	0.39 ± 0.01	3.7 ± 0.1	37.8 ± 0.4	0.99
Non-competitive	1.57 ± 0.08	6.3 ± 0.3	44.7 ± 1.1	0.98
Un-competitive	0.94 ± 0.09	8.5 ± 0.8	49.3 ± 2.5	0.95
Mixed	1.21 ± 171.1	1.9 ± 784.6	5.82 ± 1566.7	0.76

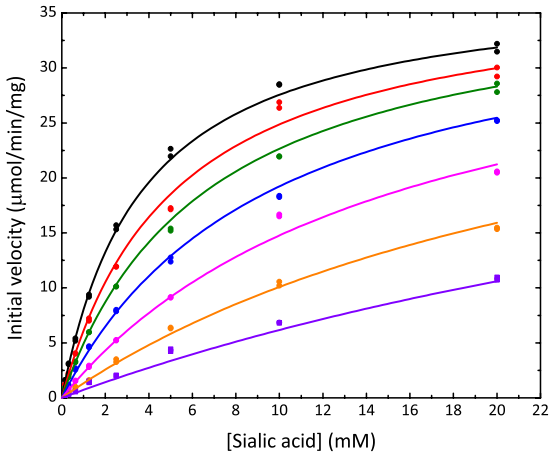


Figure 2.12. Kinetic analyses of MRSA *N*-acetylneuraminate lyase with varying concentrations of sialic acid and the sialic acid alditol mixture. Data were fitted to the competitive model of inhibition. Black corresponds to 0 mM, red corresponds to 0.156 mM, green corresponds to 0.313 mM, blue corresponds to 0.625, magenta corresponds to 1.25 mM, orange corresponds to 2.5 mM and purple corresponds to 5 mM of the sialic acid alditols.

2.2.5 Crystallisation of MRSA *N*-acetylneuraminate lyase

2.2.5.1 Crystallisation of ligand free MRSA N-*acetylneuraminate lyase*

To gain molecular insight into MRSA *N*-acetylneuraminate lyase, the crystal structure was solved. Crystallisation of ligand free MRSA *N*-acetylneuraminate lyase was successful using the JCSG+ Suite screen, which produced crystals with a needle-like morphology at 20 °C in various conditions. In-house optimisation of the JCSG+ Suite condition H7 (25% (w/v) polyethylene glycol 3350, 200 mM ammonium sulfate, 100 mM Bis Tris, pH 5.5) produced larger crystals of a similar morphology (Chapter Six, Section 6.7.1). An X-ray diffraction data set to a resolution of 1.70 Å was collected from one of these crystals.

2.2.5.2 Co-crystallisation of MRSA N-*acetylneuraminate lyase and sialic acid alditol*

For molecular insight into the inhibition of MRSA *N*-acetylneuraminate lyase by the sialic acid alditols, the structure was solved in complex with this strong inhibitor. Co-crystallisation of MRSA *N*-acetylneuraminate lyase with the sialic acid alditols was not successful in the crystallisation conditions used for ligand free MRSA *N*-acetylneuraminate lyase. Thus, co-crystallisation was performed in conditions used to grow crystals of *S. aureus N*-acetylneuraminate lyase (Timms *et al.*, 2013). The condition comprised 25% (w/v) polyethylene glycol 3350, 200 mM sodium chloride and 100 mM Tris, pH 5.5, with the addition of 5 mM of the sialic acid alditol mixture. Under these conditions, the crystals were small and rod-like in morphology. An X-ray diffraction data set to 2.33 Å was collected from one of these crystals.

2.2.6 Data processing and structure refinement

2.2.6.1 MRSA N-*acetylneuraminate lyase in a ligand free form*

X-ray diffraction data was processed and solved as described in Chapter Six, Section 6.7.2. Initially, the data was processed in the space group $P2_12_12_1$ and solved by molecular replacement using the monomer of *H. influenzae N*-acetylneuraminate lyase as the search model (PDB entry 1f6p, 54 % sequence identity) (Barbosa *et al.*, 2000). In this space group, numerous rounds of refinement failed to reduce the residual factor (R_{factor}) below 27%. This was because the data were twinned, with the twin law h;-k;-l, and an l-statistic of 0.379 (un-twinned = 0.5, perfect twin = 0.375), as implemented in *PHENIX XTRIAGE* (Adams *et al.*, 2010) (Figure 2.13). To account for twinning, data were scaled in a lower symmetry monoclinic space group ($P2_1$). Here, the asymmetric unit contains eight monomers, corresponding to two tetramers. Following refinement, identification and inclusion of the twinning law reduced the final R_{factor} to 15.5% and the free R_{factor} (R_{free}) to 20.2%. All residues were modelled into the electron density, except for the methionine residue at position one of each chain, which was too disordered. A highly conserved Tyr111 residue is in the 'disallowed' region of the Ramachandran plot in each chain. This is not unexpected as it is the case for the corresponding residue in all *N*-acetylneuraminate lyase and DHDPS structures (Izard *et al.*, 1994; Mirwaldt *et al.*, 1995; Dobson *et al.*, 2005;

Burgess *et al.*, 2008; Kefala *et al.*, 2008; Devenish & Gerrard, 2009). All relevant data collection statistics and structure and refinement statistics are provided in Table 2.2.

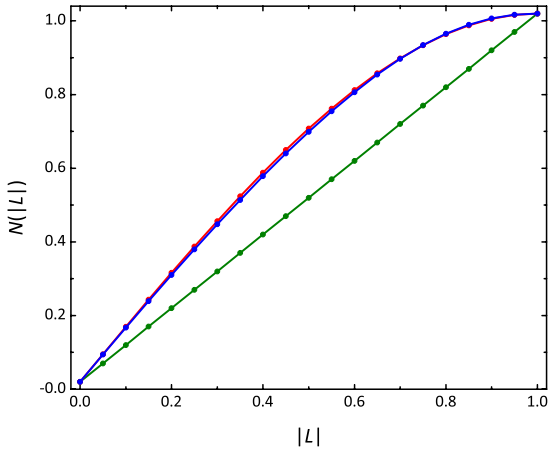


Figure 2.13. Twinning detection using the *L*-test. $|L|$ is plotted against $N(|L|)$. L is given by $(I_1-I_2)/(I_1+I_2)$, where I_1 and I_2 are the intensities of unrelated reflections and $N(|L|)$ is the cumulative probability distribution of L. The L-test suggests the acentric data could be twinned (0.379). Acentric expected and un-twinned is shown in green, acentric expected and twinned is shown in red, whereas the observed is shown in blue.

2.2.6.2 MRSA N-*acetylneuraminate lyase in complex with the sialic acid alditols*

This data was processed in the space group $P2_12_12_1$ and solved by molecular replacement using the monomer of *S. aureus N*-acetylneuraminate lyase as the search model (PDB entry 4ahp) (Timms *et al.*, 2013). Unlike my ligand free structure, this data did not appear to be twinned, which is likely to be a result of using different crystallisation conditions. In this space group, the asymmetric unit contains four monomers that correspond to one tetramer. Following iterative rounds of refinement, the R_{factor} and R_{free} were reduced to 19.6 and 26.3%, respectively. This model does not include methionine at position one, or residues Ala138 to Asn146 in each monomer. The latter residues are located within a disordered loop and could not be modelled into the structure. Similarly, this loop could not be modelled into the previously published ligand free structure of *S. aureus N*-acetylneuraminate lyase (Timms *et al.*, 2013). Again, Tyr111 of each chain is in the 'disallowed' region of the Ramachandran plot. All relevant data collection statistics and structure and refinement statistics are provided in Table 2.2.

Table 2.2. X-ray data collection and refinement statistics for ligand free, and sialic acid alditol bound MRSA *N*-acetylneuraminate lyase. Statistical values for the highest resolution shells are given in parentheses. Each monomer within the asymmetric unit has a molecular weight of 33 042.5 Da.

	Ligand free MRSA *N*-acetylneuraminate lyase	Sialic acid alditol bound MRSA *N*-acetylneuraminate lyase
data collection statistics		
wavelength (Å)	0.95369	0.95369
number of images	360	360
oscillations (°)	0.5	0.5
space group	$P2_1$	$P2_12_12_1$
unit cell parameters (Å, °)	a = 80.2, b = 108.6, c = 130.8 β = 90.14	a = 81.48, b = 109.4, c = 130.8 $\alpha = \beta = \gamma = 90$
resolution range (Å)	32.3-1.70 (1.79-1.70)	46.22-2.33 (2.41-2.33)
observed reflections	823916 (91713)	367079 (29172)
unique reflections	242175 (33054)	49980 (4071)
mean $I/\sigma(I)$	8.0 (2.6)	14.2 (2.6)
ompleteness (%)	98.5 (92.4)	99.1 (89.8)
R_{merge}†	0.101 (0.376)	0.108 (0.812)
$R_{r.i.m.}$‡	0.118 (0.457)	0.125 (0.944)
$R_{p.i.m.}$§	0.061 (0.254)	0.063 (0.478)
wilson B value (Å^2)	15.5	27.7
molecules per asymmetric unit	8	4
V_M (Å^3 Da^{-1})	2.16	2.21
solvent content (%)	43	44
structure and refinement statistics		
R_{factor} (%)	15.5	19.6
R_{free} (%)	20.2	26.3
number of atoms		
protein	18608	9080
water	942	180
ligands	40	84
root-mean-square deviation (rmsd)		
bonds (Å)	0.008	0.009
angles (°)	1.102	1.217
average B factors (Å^2)		
protein	14.62	37.50
water	16.78	35.06
ligands	19.10	42.79
Ramachandran plot, residues (%)		
favoured region	97.07	96.3
allowed region	2.59	3.3
disallowed region	0.34	0.36

R_{merge} † = $\Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$.
$R_{r.i.m}$ ‡ = $\Sigma_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$.
$R_{p.i.m.}$§ = $\Sigma_{hkl} [1/(N-1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$.

2.2.7 The structure of MRSA *N*-acetylneuraminate lyase

Consistent with the structure in solution, the crystal structure of MRSA *N*-acetylneuraminate lyase is a tetramer. The tetramer can be described as a dimer of dimers, with monomers *a* and *b* forming one dimer and monomers *c* and *d* forming the second dimer (Figure 2.14 A). This arrangement is identical to that observed for all structures of *N*-acetylneuraminate lyase enzymes studied to date, including the orientation of the active sites towards the central cavity of the tetramer. Like the *E. coli N*-acetylneuraminate lyase enzyme (Devenish & Gerrard, 2009), the intradimeric interface of MRSA *N*-acetylneuraminate lyase is significantly stronger than the interdimeric interface, with 1429 Å and 1147 Å of surface area buried, respectively. Within the monomer of MRSA *N*-acetylneuraminate lyase, consecutive β-strands are labelled *a* through *h* and the α-helices are labelled *A* through *K* with the three additional C-terminal α-helices denoted as *I*, *J*, and *K*, respectively (Figure 2.14 B). As with all $(\beta/\alpha)_8$-barrel enzymes, the active site is located in the centre of the barrel at the C-terminal end of the β-strands (Wierenga, 2001).

As described in Section 2.1.3, mutagenesis, kinetic analysis and structural analysis of *N*-acetylneuraminate lyase enzymes in complex with substrates and substrate analogues have elucidated key residues that are involved in substrate binding and catalysis (Izard *et al.*, 1994; Lawrence *et al.*, 1997; Barbosa *et al.*, 2000; Kruger *et al.*, 2001; Huynh *et al.*, 2013; Daniels *et al.*, 2014). In combination with a sequence analysis of *N*-acetylneuraminate lyase from six bacterial species (Figure 2.15), the residues that are likely to be important for the chemistry catalysed by MRSA *N*-acetylneuraminate lyase include Ser48, Ser49, Tyr111, Lys165, Thr167 Gly189, Glu191, Asp192 and Ser208. Although Ser49 is a threonine residue in other organisms, this substitution is considered conservative as these amino acids have very similar properties. The positions of these residues are highlighted in Figure 2.14 C.

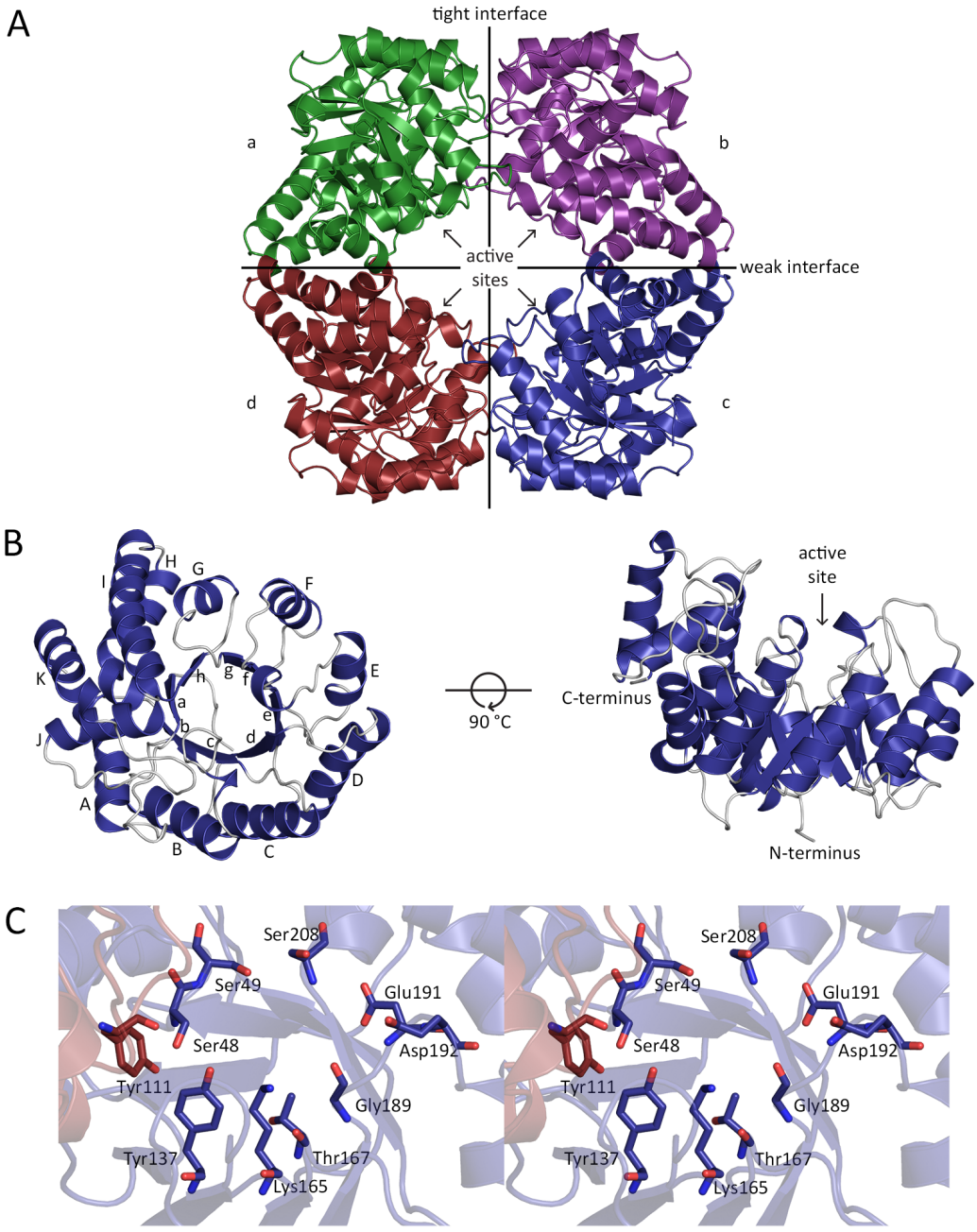


Figure 2.14. The structure of MRSA *N*-acetylneuraminate lyase. (A) The tetrameric structure of MRSA *N*-acetylneuraminate lyase. Monomers are labelled *a* through *d*. The active site of each monomer is oriented towards the central cavity of the tetramer. **(B)** The monomeric structure of MRSA *N*-acetylneuraminate lyase. The β-strands are labelled *a* through *h* and the α-helices are labelled *A* through *K*. The active site is located in the centre of the barrel at the C-terminal end of the β-strands. **(C)** A stereo view of the residues that have been implicated in substrate binding and catalysis. Tyr111 is coloured red as it is from the neighbouring subunit.

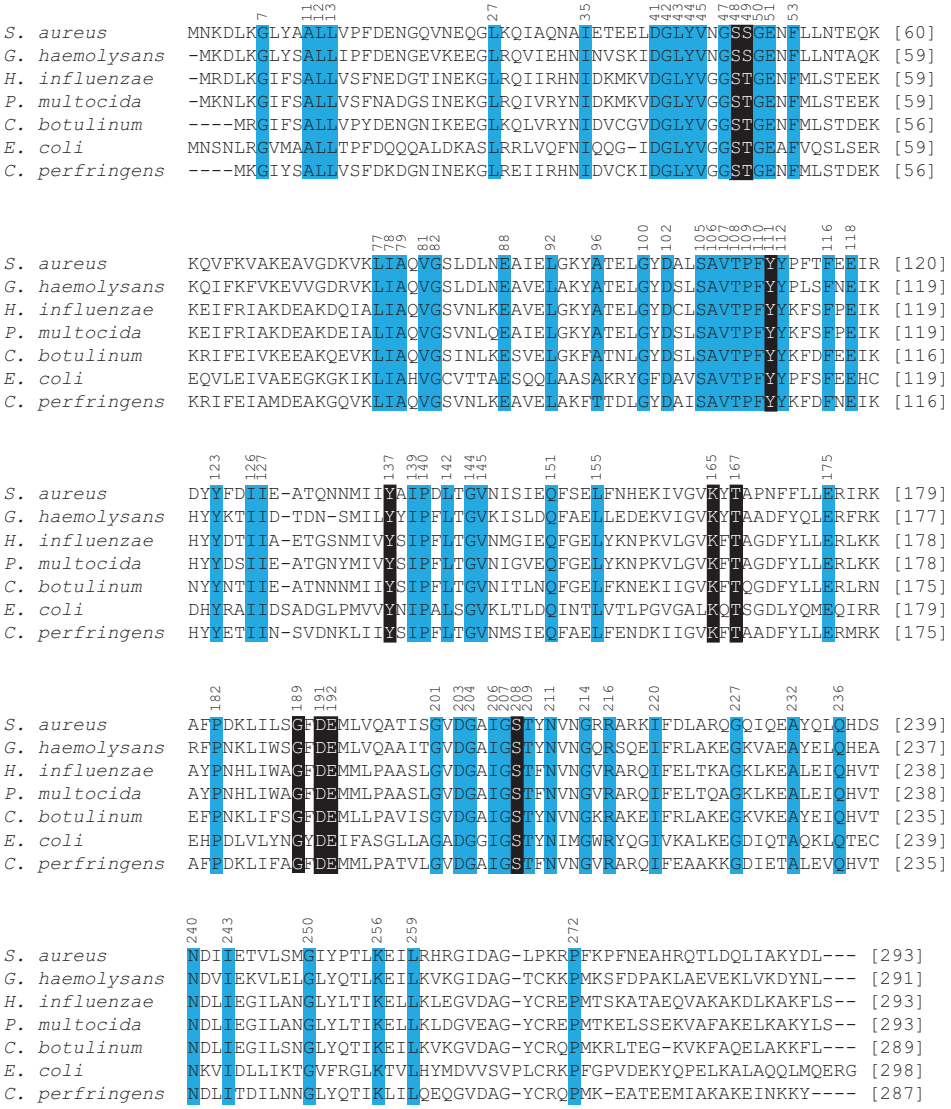


Figure 2.15. Sequence alignment of *N*-acetylneuraminate lyase enzymes from seven bacterial species. Species include *S. aureus*, *Gamella haemolysans*, *H. influenzae*, *P. multocida*, *Clostridium botulinum*, *E. coli* and *C. perfringens*. Highly conserved residues are highlighted with blue boxes, while those that have been implicated in substrate binding and catalysis are highlighted with black boxes. Conserved residues are numbered at the top of the sequence according to *S. aureus*. The alignment was generated using *ClustalW* (Larkin *et al.*, 2007) (http:// www.genome.jp/tools/clustalw/).

2.2.8 A sulfate ion residing in the active site

In the ligand free structures of *E. coli* and *H. influenzae N*-acetylneuraminate lyase (PDB entries 1nal and 1f6p, respectively), a sulfate ion was found in the active site, presumably

arising from ammonium sulfate in the crystallisation conditions (Izard *et al.*, 1994; Barbosa *et al.*, 2000). Importantly, this prevented crystallographic analysis of *N*-acetylneuraminate lyase in complex with substrates and/or substrate analogues (Izard *et al.*, 1994). Similarly, a sulfate ion was found residing in the active site of my ligand free structure of MRSA *N*-acetylneuraminate lyase. This was not unexpected as the crystallisation conditions contained 200 mM ammonium sulfate. This sulfate ion forms hydrogen bonds with the catalytically important residues Ser48, Ser49, Lys165 and Tyr137 (Figure 2.16). This perhaps explains why co-crystallisation of MRSA *N*-acetylneuraminate lyase in complex with the sialic acid alditols was unsuccessful in these conditions.

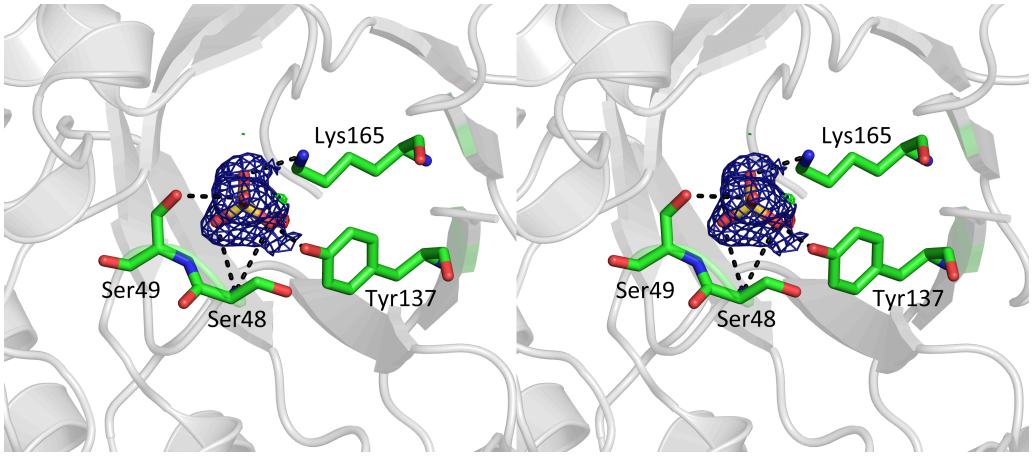


Figure 2.16. A sulfate ion in the active site of ligand free MRSA *N*-acetylneuraminate lyase. A stereo view of the sulfate ion that was modelled into electron density in the active site. The $2F_o$-F_c electron density was contoured at a sigma level of 1 (blue), while the F_o-F_c was contoured at a sigma level of 3 (green) and -3 (red). Hydrogen bonds are depicted as black dashes.

2.2.9 A comparison of the *S. aureus* and MRSA *N*-acetylneuraminate lyase structures

The structure of *S. aureus N*-acetylneuraminate lyase (PDB entry 4ahp), which has a 100% sequence identity to MRSA *N*-acetylneuraminate lyase, was solved and published by Timms *et al.* (2013) at the same time that the crystallisation and preliminary X-ray diffraction analysis of my ligand free MRSA *N*-acetylneuraminate lyase was published (North *et al.*, 2013). However, my 1.70 Å structure of MRSA *N*-acetylneuraminate lyase is

of a higher resolution than the 2.10 Å structure of PDB entry 4ahp. More importantly, PDB entry 4ahp is missing amino acid residues Ala138 to Asn146, which forms a loop between the C-terminal end of β-strand *e* and α-helix *E* in my MRSA *N*-acetylneuraminate lyase structure (Figure 2.17). Timms *et al.* (2013) propose that this loop is missing from PDB entry 4ahp because ordering is only achieved when substrate is present in the active site, allowing the formation of a Schiff base, which repositions the catalytically important Tyr137 residue for catalysis. However, the presence of this loop in my structure of MRSA *N*-acetylneuraminate lyase suggests that ordering of this loop is unlikely to be involved in substrate binding.

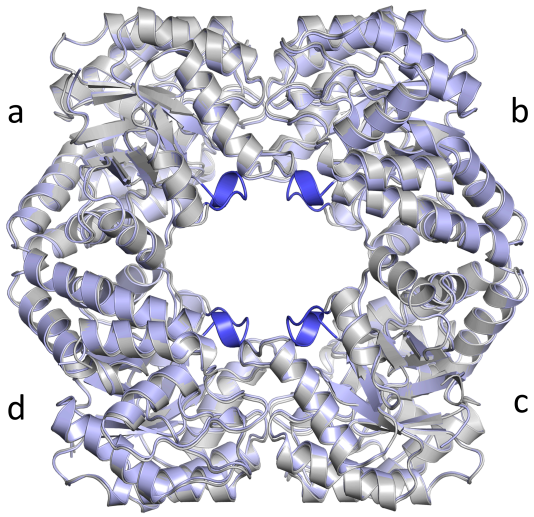


Figure 2.17. The structure of MRSA *N*-acetyleuraminate lyase and *S. aureus N*-acetylneuraminate lyase overlaid. The loop comprised of amino acid residues Ala138 to Asn146 missing from *S. aureus N*-acetylneuraminate lyase (gray) is highlighted in blue in MRSA *N*-acetylneuraminate lyase (light purple).

Furthermore, when these structures are aligned, they have an rmsd of 0.49 Å based on 1094 pairs of atoms, indicating that they are highly similar. Despite this similarity, the overall shape of these structures is slightly different, with PDB entry 4ahp being longer along the axis of the tight dimer interface, and shorter along the axis of the weak dimer interface (Figure 2.18). It is well known in the literature that crystal environments can influence protein structure (Eyal *et al.*, 2005). In alternate environments, a protein crystal will pack differently, which may result in local structural differences, rigid body motion of large structural units, or conformational changes in loops (Bertrand *et al.*, 2000; Taylor *et al.*, 2001; Eyal *et al.*, 2005). Because these structures were crystallised in different conditions,

it is reasonable to assume that both the unresolved loop formed by residues Ala138 to Asn146 in PDB entry 4ahp and the difference in overall shape between the structures is an artefact of crystal packing in alternate environments. As was shown in Figure 2.8, this would also explain why the theoretical scattering profile of PDB entry 4ahp does not fit the experimental scattering of MRSA *N*-acetylneuraminate lyase in solution as closely as my structure.

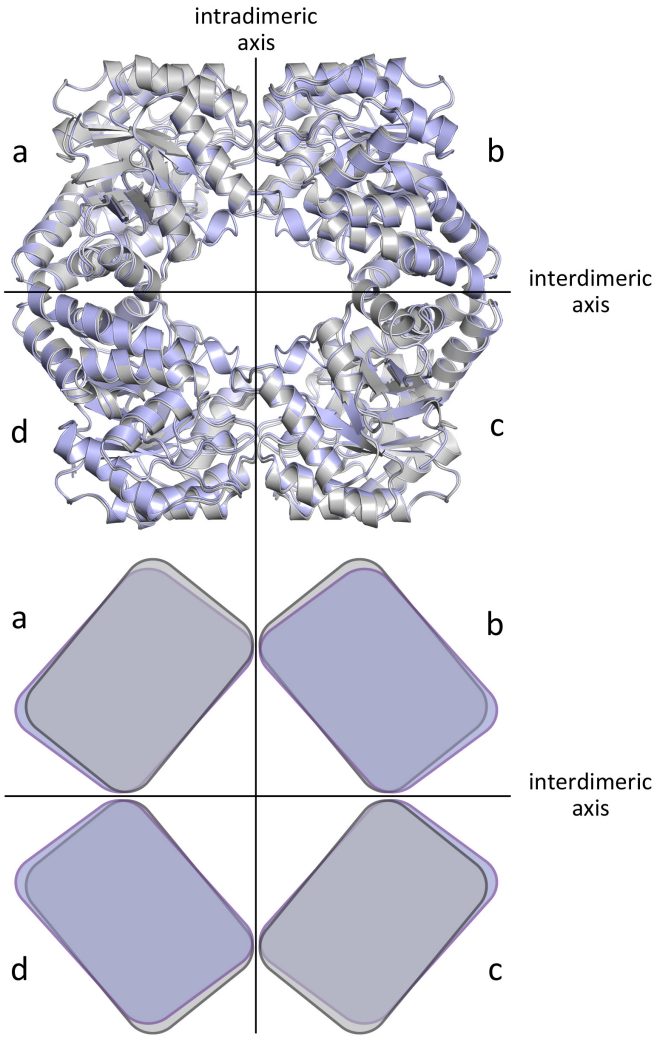


Figure 2.18. The difference in overall shape between the structures of MRSA *N*-acetylneuraminate lyase and *S. aureus N*-acetylneuraminate lyase. The crystal structure of *S. aureus N*-acetylneuraminate lyase (gray) is longer than MRSA *N*-acetylneuraminate lyase (light purple) along the axis of the intradimeric interface, and shorter along the axis of the interdimeric interface (top). This difference is emphasised in the schematic (bottom).

2.2.10 The structure of MRSA *N*-acetylneuraminate lyase in complex with a strong inhibitor

As described in Section 2.2.4, the sialic acid alditols exhibited a K_i of 0.39 mM against MRSA *N*-acetylneuraminate lyase, making it a strong inhibitor. Thus, the structure of MRSA *N*-acetylneuraminate lyase in complex with this inhibitor was sought, as this type of data is important for further optimisation and the development of antibiotics. Because crystallographic analysis of *N*-acetylneuraminate lyase in complex with substrates and/or substrate analogues is difficult in crystallisation conditions that contain sulfate (Barbosa *et al.*, 2000), co-crystallisation was performed in conditions that were without sulfate (Timms *et al.*, 2013). Following refinement, electron density was observed in all four MRSA *N*-acetylneuraminate lyase monomers in the asymmetric unit that could not be accounted for by the molecular replacement model, waters, or molecules present in the crystallisation conditions.

The unresolved loop from residue Ala138 to Asn146 in PDB entry 4ahp is also unresolved for MRSA *N*-acetylneuraminate lyase in complex with sialic acid alditol. To ensure that this is not a result of model bias, the structure was also solved by molecular replacement with my ligand free structure of MRSA *N*-acetylneuraminate lyase as the molecular replacement model, where this loop is resolved. Following numerous rounds of refinement and attempts to build in these missing residues, no electron density appeared. This further suggests that the unresolved loop in PDB entry 4ahp is likely to be an artefact of crystal packing in different environments and not a feature of substrate binding.

Although sialic acid alditol was synthesised as a racemic mixture, only the 2*R*-diastereoisomer was successfully modelled and refined into the density within each monomer (Figure 2.19 A). Sialic acid alditol binds in the active site of MRSA *N*-acetylneuraminate lyase, which is located in the centre of the $(\beta/\alpha)_8$-barrel, in a tight pocket located on the C-terminal end of the barrel (Figure 2.19 B). Importantly, this molecule has been refined into each monomer in a relatively similar conformation (Figure 2.19 C), suggesting that there is little flexibility in its binding to MRSA *N*-acetylneuraminate lyase.

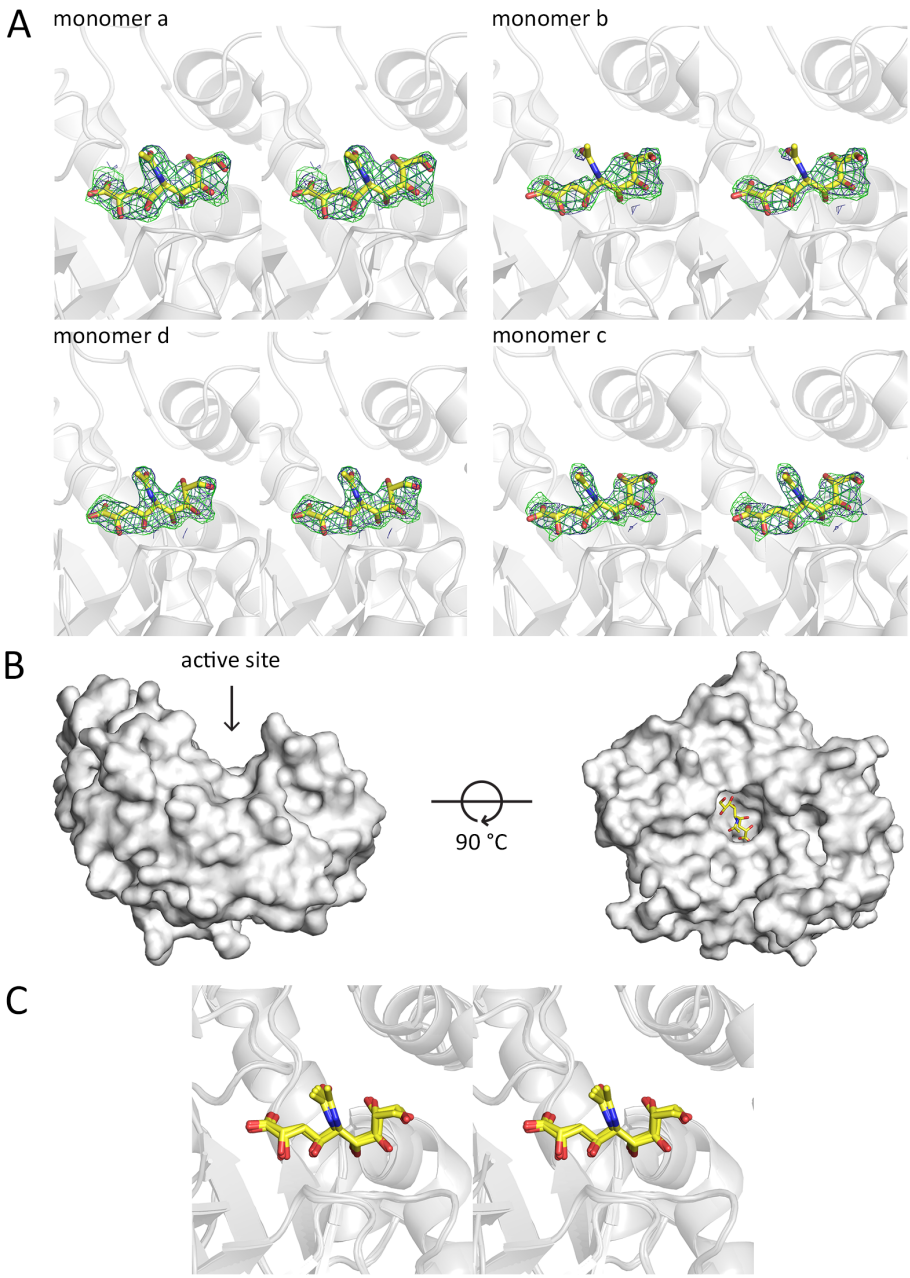


Figure 2.19. Sialic acid alditol in complex with MRSA *N*-acetylneuraminate lyase. (A) Stereo view of active site omit maps with the 2*R*-diastereoisomer of sialic acid alditol (yellow) within all four monomers in the asymmetric unit. Sialic acid alditol is depicted in an open chain form. The omit map was generated by removing the ligand from all active sites followed by refinement. The $2F_o$-F_c electron density was contoured at a sigma level of 1 (blue), while the F_o-F_c was contoured at a sigma level of 3 (green) and -3 (red). **(B)** Side view (left) and top view (right) of the MRSA *N*-acetylneuraminate lyase monomer with sialic acid alditol bound in the active site. The monomer is depicted as a surface representation and sialic acid alditol binds in a tight pocket at the center of the barrel. **(C)** Stereo view of sialic acid alditol from each monomer overlaid.

2.2.11 How inhibition is achieved by sialic acid alditol

Analysis of the binding site for the 2*R*-diastereoisomer of sialic acid alditol in MRSA *N*-acetylneuraminate lyase provides insight into how inhibition is achieved. This molecule contains a hydroxyl group at the C2 position as opposed to an aldehyde group, thus it should be incapable of Schiff base formation with Lys165. As anticipated, no covalent bond is formed between the C2 position of sialic acid alditol and Lys165 (Figure 2.20). Hydrogen bonds between Ser48 and Ser49 with the carboxylate moiety at position C1 of sialic acid alditol are formed just as they do in covalent complexes of *E. coli N*-acetylneuraminate lyase and pyruvate (Lawrence *et al.*, 1997). Additional hydrogen bond interactions between MRSA *N*-acetylneuraminate lyase and sialic acid alditol include the hydroxyl group at the C4 position with Gly189 and the hydroxyl group at the C6 position with Asp191 and Ser208; the latter of which also hydrogen bonds with the hydroxyl group at position C7 and the hydroxyl groups at position C7 and C9, which interact with Glu192. Importantly, all of these residues have been implicated in substrate/product binding for *N*-acetylneuraminate lyase enzymes (Barbosa *et al.*, 2000; Kruger *et al.*, 2001; Huynh *et al.*, 2013; Daniels *et al.*, 2014). In addition, Tyr252 hydrogen bonds with the carbonyl oxygen at position C5 of sialic acid alditol. The corresponding residue in *E. coli N*-acetylneuraminate lyase (Phe252) lies in close proximity to the C5 carbonyl of substrate/product, but cannot form a hydrogen bond (Daniels *et al.*, 2014).

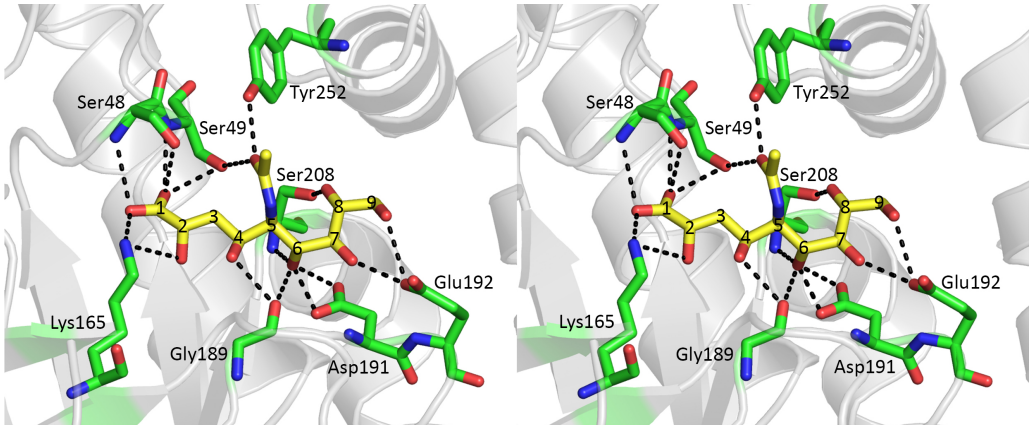


Figure 2.20. The sialic acid alditol binding site. Stereo view of the sialic acid alditol 2*R*-diastereoisomer binding site within MRSA *N*-acetylneuraminate lyase. Hydrogen bonds between the ligand and protein are indicated by black dashes.

The structure of *H. influenzae* *N*-acetylneuraminate lyase in complex with sialic acid alditol has been solved in the past. In this structure, it is also the 2*R*-diastereoisomer of sialic acid alditol that is bound in the active site. As depicted in Figure 2.21, sialic acid alditol is in a relatively similar conformation within both the MRSA and *H. influenzae* active sites. The hydroxyl groups at positions C7 and C8 have been refined in different orientations, however the active site residues implicated in binding sialic acid alditol are the same for both structures. Residue Tyr251 in *H. influenzae* *N*-acetylneuraminate lyase lies in close proximity, but does not interact, with the carbonyl oxygen at position C5 of sialic acid alditol like the corresponding residue Tyr252 in MRSA *N*-acetylneuraminate lyase, which adopts an alternate conformation.

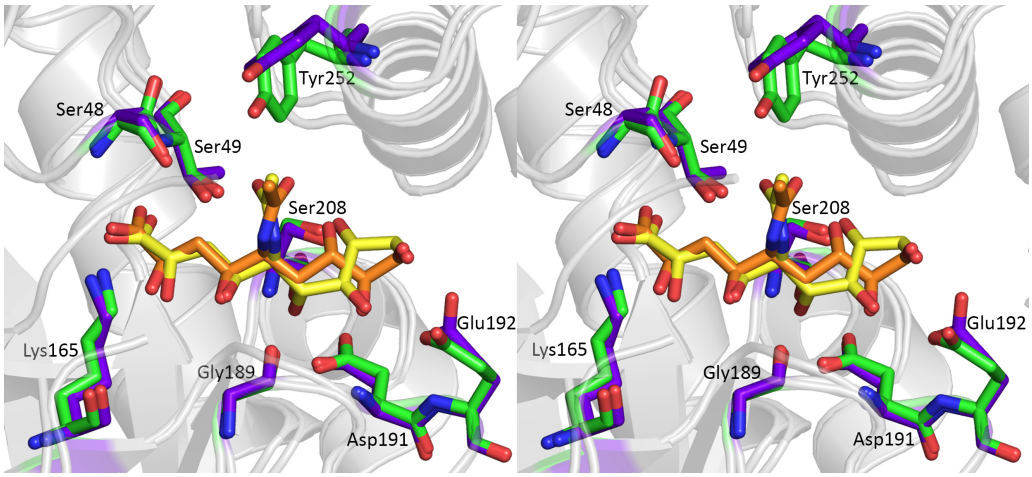


Figure 2.21. The sialic acid alditol binding site overlaid. A stereo view of the residues involved in binding the sialic acid alditol 2*R*-diastereoisomer in MRSA (green) and *H. influenzae* (purple) *N*-acetylneuraminate lyase. Sialic acid alditol modelled into MRSA *N*-acetylneuraminate lyase is yellow and in *H. influenzae* *N*-acetylneuraminate lyase it is orange.

As described, the K_i for a mixture of 2*R*- and 2*S*-diastereoisomers of sialic acid alditol with MRSA *N*-acetylneuraminate lyase is 0.39 mM, which is a magnitude better than the K_i of 4.1 mM observed for the *C. perfringens* enzyme (Deijl & Vliegenthart, 1983). This increase in binding affinity against MRSA *N*-acetylneuraminate lyase suggests that species-specific inhibition is possible, which in this case may be attributed to altered binding interactions within the active site of respective enzymes and/or subtle differences in protein dynamics in solution.

2.3 Summary

This chapter reports a detailed investigation of the structure, function and inhibition of *N*-acetylneuraminate lyase from MRSA. For the first time, solution studies confirmed that the enzyme is tetrameric in structure, consistent with *N*-acetylneuraminate lyase enzymes from other organisms. Kinetic analysis of MRSA *N*-acetylneuraminate lyase reports kinetic constants for the enzyme, which are consistent with the literature. A molecule known as sialic acid alditol, which has previously been demonstrated to inhibit the *C. perfringens N*-acetylneuraminate lyase enzyme, was synthesised and kinetic analysis established that it is a much stronger inhibitor for MRSA *N*-acetylneuraminate lyase.

Prior to this thesis, the structure of MRSA *N*-acetylneuraminate lyase had not been solved. However, the structure of *S. aureus N*-acetylneuraminate lyase was solved and published during the course of this study (PDB entry 4ahp) (Timms *et al.*, 2013). Although this is not a methicillin-resistant strain of *S. aureus*, the enzyme has a 100% sequence identity to the construct studied in this chapter. Importantly, my structure of MRSA *N*-acetylneuraminate lyase is of a higher resolution than the published *S. aureus* structure and contains new information. Specifically, my structure resolves a loop that is disordered and missing from PDB entry 4ahp.

To elucidate which residues are involved in binding sialic acid alditol in MRSA *N*-acetylneuraminate lyase, the structure was also solved in complex with this inhibitor. A small difference in the binding of this inhibitor is observed between MRSA and *H. influenzae N*-acetylneuraminate lyase. Altered binding modes and/or differences in protein dynamics within the active site of *N*-acetylneuraminate lyase enzymes may account for the species-specific inhibition observed between organisms.

2.4 References

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D, Biological Crystallography, 66*, 213-221.

Almagro-Moreno, S. & Boyd, E. F. (2009a). Insights into the evolution of sialic acid catabolism among bacteria. *BioMed Central Evolutionary Biology, 9*, 118-133.

Almagro-Moreno, S. & Boyd, E. F. (2009b). Sialic acid catabolism confers a competitive advantage to pathogenic *Vibrio cholerae* in the mouse intestine. *Infection and Immunity, 77*, 3807-3816.

Barbosa, J., Smith, B., DeGori, R., Ooi, H., Marcuccio, S., Campi, E., Jackson, W., Brossmer, R., Sommer, M. & Lawrence, M. (2000). Active site modulation in the *N*-acetylneuraminate lyase sub-family as revealed by the structure of the inhibitor-complexed *Haemophilus influenzae* enzyme. *Journal of Molecular Biology, 303*, 405-421.

Bertrand, J. A., Fanchon, E., Martin, L., Chantalat, L., Auger, G., Blanot, D., van Heijenoort, J. & Dideberg, O. (2000). "Open" structures of MurD: domain movements and structural similarities with folylpolyglutamate synthetase. *Journal of Molecular Biology, 301*, 1257-1266.

Blickling, S., Renner, C., Laber, B., Pohlenz, H. D., Holak, T. A. & Huber, R. (1997). Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochemistry, 36*, 24-33.

Buchanan, C. L., Connaris, H., Danson, M. J., Reeve, C. D. & Hough, D. W. (1999). An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. *The Biochemical Journal, 343*, 563-570.

Burgess, B. R., Dobson, R. C., Bailey, M. F., Atkinson, S. C., Griffin, M. D., Jameson, G. B., Parker, M. W., Gerrard, J. A. & Perugini, M. A. (2008). Structure and evolution of a novel dimeric enzyme from a clinically important bacterial pathogen. *Journal of Biological Chemistry, 283*, 27598-27603.

Campeotto, I., Bolt, A. H., Harman, T. A., Dennis, C., Trinh, C. H., Phillips, S. E. V., Nelson, A., Pearson, A. R. & Berry, A. (2010). Structural insights into substrate specificity in variants of *N*-acetylneuraminic acid lyase produced by directed evolution. *Journal of Molecular Biology, 404*, 56-69.

Campeotto, I., Carr, S. B., Trinh, C. H., Nelson, A. S., Berry, A., Phillips, S. E. & Pearson, A. R. (2009). Structure of an *Escherichia coli N*-acetyl-D-neuraminic acid lyase mutant, E192N, in complex with pyruvate at 1.45 Å resolution. *Acta crystallographica Section F, Structural Biology and Crystallisation Communications, 65*, 1088-1090.

Chambers, H. F. & Deleo, F. R. (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nature Reviews Microbiology, 7*, 629-641.

Chang, D. E., Smalley, D. J., Tucker, D. L., Leatham, M. P., Norris, W. E., Stevenson, S. J., Anderson, A. B., Grissom, J. E., Laux, D. C., Cohen, P. S. & Conway, T. (2004). Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 7427-7432.

Cornish-Bowden, A. (2004). *Fundamentals of enzyme kinetics*. London, Portland Press Ltd.

Daniels, A. D., Campeotto, I., van der Kamp, M. W., Bolt, A. H., Trinh, C. H., Phillips, S. E., Pearson, A. R., Nelson, A., Mulholland, A. J. & Berry, A. (2014). Reaction mechanism of *N*-acetylneuraminic acid lyase revealed by a combination of crystallography, QM/MM simulation, and mutagenesis. *American Chemical Society Chemical Biology, 9*, 1025-1032.

Deijl, C. M. & Vliegenthart, J. F. (1983). Configuration of substrate and products of *N*-acetylneuraminate pyruvate-lyase from *Clostridium perfringens*. *Biochemical and Biophysical Research Communications, 111*, 668-674.

Devenish, S. R. A. & Gerrard, J. A. (2009). The quaternary structure of *Escherichia coli N*-acetylneuraminate lyase is essential for functional expression. *Biochemical and Biophysical Research Communications, 388*, 107-111.

Dobson, R. C., Griffin, M. D., Jameson, G. B. & Gerrard, J. A. (2005). The crystal structures of native and (S)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. *Acta Crystallographica Section D, Biological Crystallography, 61*, 1116-1124.

Dobson, R. C., Valegard, K. & Gerrard, J. A. (2004). The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: further evidence for a catalytic triad. *Journal of Molecular Biology, 338*, 329-339.

Eaton, R. W. (1994). Organisation and evolution of naphthalene catabolic pathways: sequence of the DNA encoding 2-hydroxychromene-2-carboxylate isomerase and *trans*-*o*-hydroxybenzylidenepyruvate hydratase-aldolase from the NAH7 plasmid. *Journal of Bacteriology, 176*, 7757-7762.

Eyal, E., Gerzon, S., Potapov, V., Edelman, M. & Sobolev, V. (2005). The limit of accuracy of protein modeling: influence of crystal packing on protein structure. *Journal of Molecular Biology, 351*, 431-442.

Griffin, M. D., Dobson, R. C., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. & Gerrard, J. A. (2008). Evolution of quaternary structure in a homotetrameric enzyme. *Journal of Molecular Biology, 380*, 691-703.

Gross, H. J. & Brossmer, R. (1988). Inhibition of *N*-acetylneuraminate lyase by *N*-acetyl-4-oxo-D-neuraminic acid. *Federation of European Biochemical Societies Letters, 232*, 145-147.

Grundmann, H., Aires-de-Sousa, M., Boyce, J. & Tiemersma, E. (2006). Emergence and resurgence of meticillin-resistant *Staphylococcus aureus* as a public-health threat. *Lancet, 368*, 874-885.

Hagedorn, H. W. & Brossmer, R. (1986). Synthesis and biological properties of *N*-acetyl-4-deoxy-D-neuraminic acid. *Helvetica Chimica Acta, 69*, 2127–2132.

Huynh, N., Aye, A., Li, Y., Yu, H., Cao, H., Tiwari, V. K., Shin, D. W., Chen, X. & Fisher, A. J. (2013). Structural basis for substrate specificity and mechanism of *N*-acetyl-D-neuraminic acid lyase from *Pasteurella multocida*. *Biochemistry, 52*, 8570-8579.

Iwabuchi, T. & Harayama, S. (1998). Biochemical and genetic characterisation of *trans*-2'-carboxybenzalpyruvate hydratase-aldolase from a phenanthrene-degrading *Nocardioides* strain. *Journal of Bacteriology, 180*, 945-949.

Izard, T., Lawrence, M. C., Malby, R. L., Lilley, G. G. & Colman, P. M. (1994). The three-dimensional structure of *N*-acetylneuraminate lyase from *Escherichia coli*. *Structure, 2*, 361-369.

Jeffcoat, R., Hassall, H. & Dagley, S. (1969a). The metabolism of D-glucarate by *Pseudomonas acidovorans*. *The Biochemical Journal, 115*, 969-976.

Jeffcoat, R., Hassall, H. & Dagley, S. (1969b). Purification and properties of D-4-deoxy-5-oxoglucarate hydrolyase (decarboxylating). *The Biochemical Journal, 115*, 977-983.

Jeong, H. G., Oh, M. H., Kim, B. S., Lee, M. Y., Han, H. J. & Choi, S. H. (2009). The capability of catabolic utilisation of *N*-acetylneuraminic acid, a sialic acid, is essential for *Vibrio vulnificus* pathogenesis. *Infection and Immunity, 77*, 3209-3217.

Joerger, A. C., Mayer, S. & Fersht, A. R. (2003). Mimicking natural evolution *in vitro*: an *N*-acetylneuraminate lyase mutant with an increased dihydrodipicolinate synthase activity. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 5694-5699.

Kefala, G., Evans, G. L., Griffin, M. D., Devenish, S. R., Pearce, F. G., Perugini, M. A., Gerrard, J. A., Weiss, M. S. & Dobson, R. C. (2008). Crystal structure and kinetic study of dihydrodipicolinate synthase from *Mycobacterium tuberculosis*. *The Biochemical Journal, 411*, 351-360.

Kruger, D., Schauer, R. & Traving, C. (2001). Characterisation and mutagenesis of the recombinant *N*-acetylneuraminate lyase from *Clostridium perfringens*: insights into the reaction mechanism. *European Journal of Biochemistry, 268*, 3831-3839.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*, 2947-2948.

Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Cambridge, The Royal Society of Chemistry.

Lawrence, M., Barbosa, J., Smith, B., Hall, N., Pilling, P., Ooi, H. & Marcuccio, S. (1997). Structure and mechanism of a sub-family of enzymes related to *N*-acetylneuraminate lyase. *Journal of Molecular Biology, 266*, 381-399.

Lebowitz, J., Lewis, M. S. & Schuck, P. (2002). Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Science, 11*, 2067-2079.

Li, Y., Yu, H., Cao, H., Lau, K., Muthana, S., Tiwari, V. K., Son, B. & Chen, X. (2008). *Pasteurella multocida* sialic acid aldolase: a promising biocatalyst. *Applied Microbiology and Biotechnology, 79*, 963-970.

Lilley, G. G., Itzstein, M. & Ivancic, N. (1992). High-level production and purification of *Escherichia coli N*-Acetylneuraminic acid aldolase (EC 4.1. 3.3). *Protein expression and purification, 3*, 434-440.

Martinez, J., Steenbergen, S. & Vimr, E. (1995). Derived structure of the putative sialic acid transporter from *Escherichia coli* predicts a novel sugar permease domain. *Journal of Bacteriology, 177*, 6005-6010.

Mirwaldt, C., Korndorfer, I. & Huber, R. (1995). The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. *Journal of Molecular Biology, 246*, 227-239.

North, R. A., Kessans, S. A., Atkinson, S. C., Suzuki, H., Watson, A. J., Burgess, B. R., Angley, L. M., Hudson, A. O., Varsani, A., Griffin, M. D., Fairbanks, A. J. & Dobson, R. C. (2013). Cloning, expression, purification, crystallisation and preliminary X-ray diffraction studies of *N*-acetylneuraminate lyase from methicillin-resistant *Staphylococcus aureus*. *Acta crystallographica Section F, Structural Biology and Crystallisation Communications, 69*, 306-312.

Olson, M. E., King, J. M., Yahr, T. L. & Horswill, A. R. (2013). Sialic acid catabolism in *Staphylococcus aureus*. *Journal of Bacteriology, 195*, 1779-1788.

Ooi, H. C., Marcuccio, S. M. & Jackson, W. R. (1999). An alternative synthesis of *N*-acetyl-4-deoxyneuraminic acid. *Australian Journal of Chemistry, 52*, 937-940.

Ooi, H. C., Marcuccio, S. M. & Jackson, W. R. (2000). A new preparation of the diastereoisomeric *N*-acetylneuraminic alditols. *Australian Journal of Chemistry, 53*, 171-174.

Petoukhov, M. V. & Svergun, D. I. (2007). Analysis of X-ray and neutron scattering from biomacromolecular solutions. *Current opinion in structural biology, 17*, 562-571.

Pezzicoli, A., Ruggiero, P., Amerighi, F., Telford, J. L. & Soriani, M. (2012). Exogenous sialic acid transport contributes to group B streptococcus infection of mucosal surfaces. *Journal of Infectious Diseases, 206*, 924-931.

Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics, 40*, 191-285.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical Journal, 78*, 1606-1619.

Scopes, R. K. (2002). Enzyme activity and assays. *Encyclopedia of Life Sciences*, 1-6.

Severi, E., Hood, D. W. & Thomas, G. H. (2007). Sialic acid utilisation by bacterial pathogens. *Microbiology, 153*, 2817-2822.

Severi, E., Randle, G., Kivlin, P., Whitfield, K., Young, R., Moxon, R., Kelly, D., Hood, D. & Thomas, G. H. (2005). Sialic acid transport in *Haemophilus influenzae* is essential for lipopolysaccharide sialylation and serum resistance and is dependent on a novel tripartite ATP-independent periplasmic transporter. *Molecular Microbiology, 58*, 1173-1185.

Svergun, D., Barberato, C. & Koch, M. H. J. (1995). CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography, 28*, 768-773.

Svergun, D. I. (1992). Determination of the regularisation parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography, 25*, 495-503.

Taylor, P., Dornan, J., Carrello, A., Minchin, R. F., Ratajczak, T. & Walkinshaw, M. D. (2001). Two structures of cyclophilin 40: folding and fidelity in the TPR domains. *Structure, 9*, 431-438.

Timms, N., Windle, C. L., Polyakova, A., Ault, J. R., Trinh, C. H., Pearson, A. R., Nelson, A. & Berry, A. (2013). Structural insights into the recovery of aldolase activity in *N*-acetylneuraminic acid lyase by replacement of the catalytically active lysine with gamma-thialysine by using a chemical mutagenesis strategy. *Chembiochem: a European Journal of Chemical Biology, 14*, 474-481.

Vimr, E. R., Kalivoda, K. A., Deszo, E. L. & Steenbergen, S. M. (2004). Diversity of microbial sialic acid metabolism. *Microbiology and Molecular Biology Reviews, 68*, 132-153.

Vimr, E. R. & Troy, F. A. (1985). Identification of an inducible catabolic system for sialic acids (*nan*) in *Escherichia coli*. *Journal of Bacteriology, 164*, 845-853.

Vogel-Scheel, J., Alpert, C., Engst, W., Loh, G. & Blaut, M. (2010). Requirement of purine and pyrimidine synthesis for colonisation of the mouse intestine by *Escherichia coli*. *Applied and Environmental Microbiology, 76*, 5181-5187.

von Itzstein, M. (2007). The war against influenza: discovery and development of sialidase inhibitors. *Nature Reviews Drug Discovery, 6*, 967-974.

Wierenga, R. K. (2001). The TIM-barrel fold: a versatile framework for efficient enzymes. *Federation of European Biochemical Societies Letters, 492*, 193-198.

Zbiral, E., Kleineidam, R. G., Schreiner, E., Hartmann, M., Christian, R. & Schauer, R. (1992). Elucidation of the topological parameters of *N*-acetylneuraminic acid and some analogues involved in their interaction with the *N*-acetylneuraminate lyase from *Clostridium perfringens*. *The Biochemical Journal, 282*, 511-516.

Chapter Three

The structure and mechanism of *N*-acetylmannosamine-6-phosphate 2-epimerase from methicillin-resistant *Staphylococcus aureus*

Part of the work presented in this chapter has been published in:

North, R. A., Kessans, S. A., Suzuki, H., Watson, A. J. A. Griffin, M. D. W., Fairbanks, A. J. & Dobson, R. C. J. (2014). Cloning, expression, purification, crystallisation and preliminary X-ray diffraction analysis of *N*-acetylmannosamine-6-phosphate 2-epimerase from methicillin resistant *Staphylococcus aureus*. *Acta Crystallographica Section F, Structural Biology and Crystallisation Communications*, *70*, 643-649.

3.1 Introduction

3.1.1 Epimerase enzymes

Chirality is a ubiquitous feature of proteins that allows enzymes their ability to catalyse reactions with an extremely high degree of stereoselectivity (Tanner, 2002). Epimerases are enzymes capable of catalysing selective inversions of stereochemistry in biological molecules that have multiple chiral centres. To catalyse an inversion of stereochemistry at a chiral centre, an epimerase must break a bond from one face and return it to the other face in a non-stereoselective manner (Allard *et al.*, 2001; Tanner, 2002). Carbohydrate

molecules in particular have many chiral centres and almost every stereocentre is epimerised by at least one enzyme. Inversions of carbohydrate stereochemistry by epimerase enzymes can yield a diverse range of structures that cannot be generated in nature from non-carbohydrate precursors (Samuel & Tanner, 2002).

3.1.2 Epimerase enzymes implicated in sialic acid utilisation

To date, three different types of epimerase enzymes have been implicated in pathways that utilise sialic acid in both eukaryotes and prokaryotes; UDP-*N*-acetylglucosamine 2-epimerase, *N*-acetylglucosamine 2-epimerase and *N*-acetylmannosamine-6-phosphate 2-epimerase (Ferrero *et al.*, 2007). In eukaryotes, UDP-*N*-acetylglucosamine 2-epimerase and *N*-acetylglucosamine 2-epimerase synthesise and degrade sialic acid, respectively (Hinderlich *et al.*, 1997; Stasche *et al.*, 1997; Luchansky *et al.*, 2003). Interestingly, the *N*-acetylglucosamine 2-epimerase that degrades sialic acid in eukaryotes, has also been identified in some photosynthetic cyanobacteria (Lee *et al.*, 2007; Sola-Carvajal *et al.*, 2012). The mammalian UDP-*N*-acetylglucosamine 2-epimerase exists as a bifunctional enzyme that also possesses *N*-acetylmannosamine kinase activity. Thus, as the stereochemistry of UDP-*N*-acetylglucosamine is epimerised, the *N*-acetylmannosamine kinase portion of the enzyme phosphorylates *N*-acetylmannosamine into *N*-acetylmannosamine-6-phosphate (Hinderlich *et al.*, 1997; Stasche *et al.*, 1997; Tanner, 2005).

In prokaryotes, an evolutionarily and mechanistically related UDP-*N*-acetylglucosamine 2-epimerase is involved in the synthesis of sialic acid, but this enzyme is not bifunctional (Tanner, 2005). For the degradation of sialic acid, an *N*-acetylmannosamine-6-phosphate 2-epimerase has been identified (Vann *et al.*, 1997; Plumbridge & Vimr, 1999; Walters *et al.*, 1999; Ringenberg *et al.*, 2003; Murkin *et al.*, 2004), which differs from the equivalent eukaryotic enzyme in that the substrate/product is not phosphorylated. The remainder of this chapter is focused on the *N*-acetylmannosamine-6-phosphate 2-epimerase enzyme, responsible for the inversion of stereochemistry at the C2 position of *N*-acetylmannosamine-6-phosphate, yielding *N*-acetylglucosamine-6-phosphate (Figure 3.1).

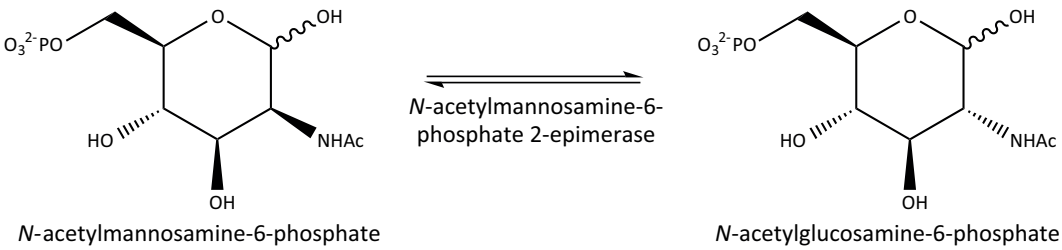


Figure 3.1. The reaction catalysed by *N*-acetylmannosamine-6-phosphate 2-epimerase. The stereochemistry of *N*-acetylmannosamine-6-phosphate is inverted at the C2 position by *N*-acetylmannosamine-6-phosphate 2-epimerase to yield *N*-acetylglucosamine-6-phosphate.

3.1.3 The catalytic mechanism of *N*-acetylmannosamine-6-phosphate 2-epimerase

To date, several mechanistic strategies have been classified for carbohydrate epimerisation, including deprotonation/reprotonation, formation of a transient keto intermediate, carbon-carbon bond cleavage, nucleotide elimination and mutarotation (Figure 3.2) (Allard *et al.*, 2001; Samuel & Tanner, 2002). Prior to this thesis, the reaction mechanism of *N*-acetylmannosamine-6-phosphate 2-epimerase had not been elucidated. However, recent crystal structures and proton nuclear magnetic resonance spectroscopy (^{1}H NMR) kinetic analysis of *Clostridium perfringens N*-acetylmannosamine-6-phosphate 2-epimerase have provided insight into the residues involved in substrate binding and the chemistry catalysed by this enzyme (Pelissier *et al.*, 2014). They propose that epimerisation by *N*-acetylmannosamine-6-phosphate 2-epimerase occurs *via* a deprotonation/reprotonation mechanism, which is the most commonly employed catalytic strategy of carbohydrate epimerases (Tanner, 2002).

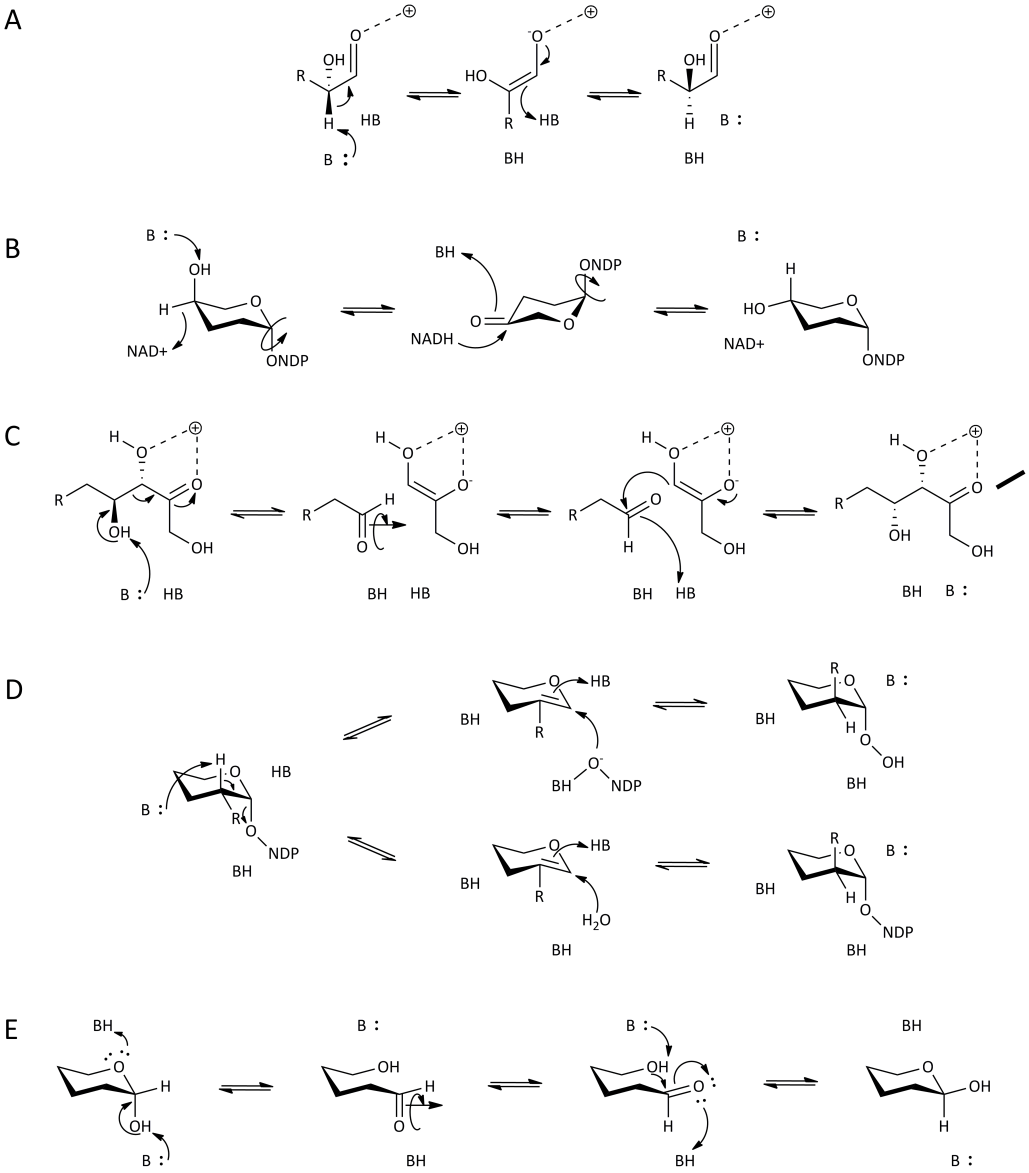


Figure 3.2. Five possible reaction mechanisms classified for carbohydrate epimerases. (A) Deprotonation/reprotonation. **(B)** Formation of a transient keto intermediate. **(C)** Carbon-carbon bond cleavage. **(D)** Nucleotide elimination. **(E)** Mutarotation. In the deprotonation/reprotonation mechanism, the positive charge can be either a divalent metal ion or a positively charged amino acid residue. In the carbon-carbon bond cleavage mechanism, the positive charge is a divalent metal ion. NAD+ and NADH respectively denote the oxidised and reduced form of nicotinamide adenine dinucleotide. NDP denotes nucleoside diphosphate [adapted from van Overtveldt *et al.* (2015)].

Deprotonation/reprotonation involves cleavage of a C-H bond at an "activated" stereogenic centre. Activated stereogenic centres are adjacent to a carbonyl group, such as an aldehyde,

ketone or carboxyl (Tanner, 2002; van Overtveldt *et al.*, 2015). The presence of this group lowers the pK_a of the stereogenic centre to less than 30, allowing the proton to be abstracted. Deprotonation forms a resonance-stabilised carbanion and then a non-stereospecific reprotonation generates the epimer (Samuel & Tanner, 2002).

For *C. perfringens* *N*-acetylmannosamine-6-phosphate 2-epimerase (Figure 3.3), it is proposed that a proton is abstracted from the C2 position of *N*-acetylmannosamine-6-phosphate by Lys66, which acts as a Brønsted base (Pelissier *et al.*, 2014). Importantly, the C2 position of *N*-acetylmannosamine-6-phosphate is adjacent to an aldehyde at the C1 position, activating the stereogenic centre for deprotonation. Following abstraction of the proton, Pelissier *et al.* (2014) suggest that a negatively charged enolate intermediate is stabilised by Arg43, which forms a salt bridge with Asp126. Reprotonation by Lys66 then generates the C2 epimer, *N*-acetylglucosamine-6-phosphate, although this requires that the substrate will rotate in the active site to allow Lys66 to access the opposing face. There is no experimental evidence that this occurs.

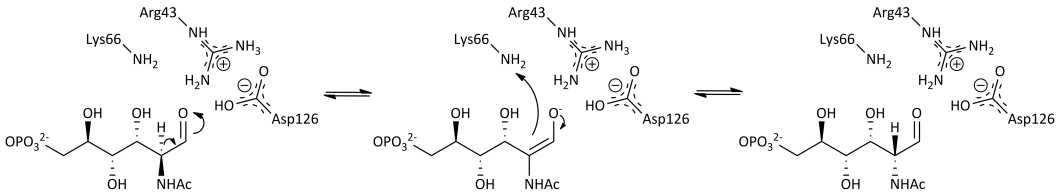


Figure 3.3. The proposed deprotonation/reprotonation catalytic mechanism of *C. perfringens* *N*-acetylmannosamine-6-phosphate 2-epimerase. A proton is abstracted by Lys66 from an activated stereogenic centre at the C2 position of *N*-acetylmannosamine-6-phosphate. The negatively charged enolate intermediate is stabilised by Arg43, forming a salt bridge with Asp126. Reprotonation by Lys66 generates the epimer, *N*-acetylglucosamine-6-phosphate [adapted from Pelissier *et al.* (2014)].

3.1.4 Kinetic analysis of *N*-acetylmannosamine-6-phosphate 2-epimerase

The current method in the literature to assay *N*-acetylglucosamine-6-phosphate 2-epimerase catalysis is by ^{1}H NMR (Pelissier *et al.*, 2014). However, this is a time and resource intensive procedure (Exnowitz *et al.*, 2012). I developed a new multi-enzyme coupled assay to monitor *N*-acetylmannosamine-6-phosphate 2-epimerase activity (Section

3.2.4). This uses the next enzymes in the sialic acid degradation pathway, *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase, along with phosphoglucoisomerase and glucose-6-phosphate dehydrogenase.

Although enzyme reactions can be followed using a variety of techniques, the most convenient way to measure the conversion of substrate to product over time is by following the change in absorbance at a particular wavelength with a spectrophotometer. For this to be possible, the change in absorbance must be induced by a chromophore that is proportional to the amount of substrate/product (Cornish-Bowden, 2004). Often, however, neither the substrate nor product absorbs light at a wavelength compatible with this technique. In this case the reaction may be able to be coupled to another enzyme reaction, which can be monitored at a particular wavelength by following the change in absorbance over time.

The reaction catalysed by *N*-acetylmannosamine-6-phosphate 2-epimerase involves the inversion of *N*-acetylmannosamine-6-phosphate and *N*-acetylglucosamine-6-phosphate 2-epimerase, neither of which can be monitored spectrophotometrically. Pelissier *et al.* (2014) used ^{1}H NMR to follow the conversion of *N*-acetylmannosamine-6-phosphate to *N*-acetylglucosamine-6-phosphate 2-epimerase, which is the direction that the equilibrium lies (Ringenberg *et al.*, 2003). Although it is not possible to directly follow the reaction catalysed by *N*-acetylmannosamine-6-phosphate 2-epimerase spectrophotometrically, it is possible to couple this reaction to other enzyme reactions until a change in absorbance can be achieved by catalysis.

3.1.5 *N*-Acetylmannosamine-6-phosphate 2-epimerase is an antibiotic drug target

As described in Chapter One, various studies have suggested that the sialic acid degradation pathway is a candidate for antibiotic drug development (Chang *et al.*, 2004; Almagro-Moreno & Boyd, 2009; Jeong *et al.*, 2009; Pezzicoli *et al.*, 2012; Olson *et al.*, 2013). Deletion of the *nan*E gene coding for *N*-acetylmannosamine-6-phosphate 2-epimerase renders *Staphylococcus aureus* (Olson *et al.*, 2013), *Escherichia coli* and *Vibrio cholerae* (Almagro-Moreno & Boyd, 2009) incapable of growing on sialic acid as a sole

carbon source *in vitro*. Although there are a number of epimerase enzymes involved in sialic acid biosynthetic or degradation pathways throughout eukaryotes and prokaryotes (Section 3.1.3), bacterial *N*-acetylmannosamine-6-phosphate 2-epimerase is highly conserved among gram-positive and gram-negative bacteria (North *et al.*, 2014), but does not display significant homology in the structure and catalytic mechanism to that of mammalian epimerases (Tanner, 2005; Pelissier *et al.*, 2014). Thus, similar to *N*-acetylneuraminate lyase (Chapter Two), *N*-acetylmannosamine-6-phosphate 2-epimerase can be considered as a potential target for the development of novel antibiotic drugs.

3.1.6 Overview of this chapter

This chapter will address the structure and catalytic mechanism of *N*-acetylmannosamine-6-phosphate 2-epimerase from methicillin resistant *S. aureus* (MRSA), a pathogen that causes significant disease globally (McDonald *et al.*, 1981; Boyce, 1992; Panlilio *et al.*, 1992; Grundmann *et al.*, 2006; Chambers & Deleo, 2009). A specifically designed assay for kinetic analysis was developed for *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes. The structure of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase was solved and in combination with a sequence analysis, residues that have potential to be involved in catalysis were identified and tested by kinetic analysis. Although structures and a proposed catalytic mechanism for *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase have been published during the course of this study, a critical analysis of these structures suggests that deprotonation/reprotonation by a conserved lysine residue is unlikely. Instead, I propose that catalysis occurs *via* the formation of a Schiff base or a proton displacement mechanism mediated by the substrate. Overall, this chapter presents a solution based characterisation of the wild type enzyme and a kinetic and structural analysis of the wild type and mutant enzymes. Because *N*-acetylmannosamine-6-phosphate 2-epimerase is a potential target for the development of antibiotic drugs, this type of information is critical for the development of first generation inhibitors.

3.2 Results and discussion

3.2.1 Cloning of *N*-acetylmannosamine-6-phosphate 2-epimerase, *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase from MRSA

The *nan*E gene encoding *N*-acetylmannosamine-6-phosphate 2-epimerase from MRSA USA300 (Appendix) was commercially supplied with *Sac*I and *Hind*III restriction sites in a cloning plasmid designated pUC57-Kan. The *nan*E gene was excised from pUC57-Kan and sub cloned into the pET30ΔSE expression plasmid (Suzuki *et al.*, 2014) to produce a construct named pET30ΔSE/*nan*E-His (Chapter Six, Sections 6.4.6-6.4.8). Initially, *Sac*I and *Hind*III restriction sites were selected so that the overexpressed protein would contain an N-terminal 6 × Histidine tag (His-tag) located within pET30ΔSE that can be cleaved off with thrombin. However, the His-tag interfered with the purified protein and was deemed inappropriate for this construct (Section 3.2.2). To remove it, the polymerase chain reaction (PCR) was used to incorporate *Nde*I and *Hind*III restriction sites into the 5' and 3' ends of the *nanE* gene in pUC57-Kan (Chapter Six, Section 6.4.5). To do this, primer pair NanE-forward and NanE-reverse (Appendix) were used. The resulting PCR product was sub cloned into pCR2.1-TOPO (Chapter Six, Section 6.4.5) and then into pET30ΔSE to create pET30ΔSE/*nan*E (Chapter Six, Sections 6.4.6-6.4.8). Deoxyribonucleic acid (DNA) sequencing was used to confirm the identity of pET30ΔSE/*nan*E-His and pET30ΔSE/*nan*E.

Mutated variants of the *nan*E gene encoding *N*-acetylmannosamine-6-phosphate 2-epimerase from MRSA USA300 were commercially supplied in a pET30a(+) expression plasmid. Each variant contained a single amino acid change and was named according to the position of the mutation. The rationale behind each of these single amino acid substitutions is discussed in more detail later (Section 3.2.8). These mutants are referred to as Gln11Ala, Gln11Ser, Glu16Ala, Glu16His, Arg40Ala, Lys63Ala, Lys63Glu, Asp124Ala, Asp124Gln and Arg208Ala throughout this thesis.

Development of a multi-enzyme coupled assay for MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase (Section 3.2.4) required the two subsequent enzymes of the sialic acid degradation pathway (*N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase). The *nag*A gene encoding *N*-acetylglucosamine-6-phosphate deacetylase and the *nag*B gene encoding glucosamine-6-phosphate deaminase, both from MRSA USA300 (Appendix), were commercially supplied with *Nde*I and *Hind*III restriction sites in pUC57-Kan. The genes were excised from pUC57-Kan and sub cloned into pET30ΔSE to produce constructs named pET30ΔSE/*nag*A and pET30ΔSE/*nag*B, respectively (Chapter Six, Sections 6.4.6-6.4.8). DNA sequencing was used to confirm the identities of these constructs.

3.2.2 Purification of *N*-acetylmannosamine-6-phosphate 2-epimerase, *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase from MRSA

All constructs to be overexpressed were transformed into *E. coli* BL21 (DE3) bacterial cells, cultured and overexpressed in Luria Bertani medium, harvested and lysed (Chapter Six, Sections 6.6.1-6.6.3). Utilisation of the *Sac*I and *Hind*III restriction sites in the pET30ΔSE plasmid enabled the incorporation of a His-tag at the N-terminus of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase (pET30ΔSE/*nan*E-His). Following purification by immobilised metal affinity chromatography, cleavage of the His-tag with thrombin proved to be unsuccessful. This suggests that the His-tag may have been interacting with the protein in a way that made the thrombin cleavage site inaccessible for the protease. Thus, utilisation of a His-tag for purification of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase was not appropriate.

The modified gene construct (pET30ΔSE/*nan*E) was expressed without a His-tag. Protein expressed from this construct was subjected to purification by anion exchange chromatography, hydrophobic interaction chromatography and size exclusion chromatography. The increasing purity of the wild type enzyme from each purification procedure is shown in Figure 3.4 A. The final purity was visually estimated to be at least 95%, observed by the single band at 24.5 kilodalton (kDa) following size exclusion chromatography (Figure 3.4 A, lane 5). Following protein purification, the mass of MRSA

N-acetylmannosamine-6-phosphate 2-epimerase was confirmed by electrospray ionisation mass spectrometry.

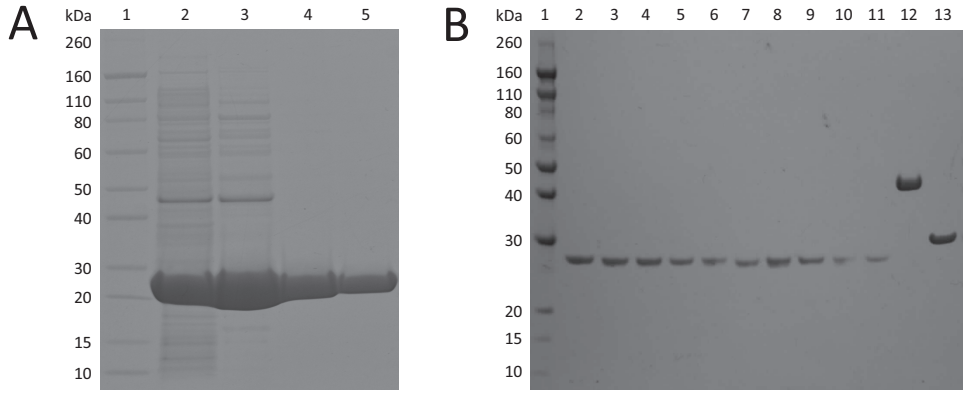


Figure 3.4. Purification of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, MRSA *N*-acetylglucosamine-6-phosphate deacetylase and MRSA glucosamine-6-phosphate deaminase. (A) Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis showing the increasing purity of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase at each purification step. Lane 1, protein ladder (kDa); lane 2, crude; lane 3, pooled fractions from anion-exchange chromatography; lane 4, pooled fractions from hydrophobic interaction chromatography; lane 5, post-size-exclusion chromatography. **(B)** SDS-PAGE analysis showing the final purity of the mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes, MRSA *N*-acetylglucosamine-6-phosphate deacetylase and MRSA glucosamine-6-phosphate deaminase. Lane 1, protein ladder (kDa); lane 2, Gln11Ala; lane 3, Gln11Ser; lane 4, Glu16Ala; lane 5, Glu16His; lane 6, Arg40Ala; lane 7, Lys63Ala; lane 8, Lys63Glu; lane 9, Asp124Ala; lane 10, Asp124Gln; lane 11, Arg208Ala; lane 12, MRSA *N*-acetylglucosamine-6-phosphate deacetylase; lane 13, MRSA glucosamine-6-phosphate deaminase.

The mutated variants of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, the MRSA *N*-acetylglucosamine-6-phosphate deacetylase and the MRSA glucosamine-6-phosphate deaminase were purified using anion exchange chromatography, hydrophobic interaction chromatography and size exclusion chromatography. The final purity of each enzyme was visually estimated to be at least 95%, observed by the single bands following size exclusion chromatography (Figure 3.4 B). The mutated variants of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase are all approximately 24.5 kDa in size (Figure 3.4 B, lanes 2 through to 11), whereas MRSA *N*-acetylglucosamine-6-phosphate deacetylase (Figure 3.4 B, lane 12) and MRSA glucosamine-6-phosphate deaminase (Figure 3.4 B, lane 13) are 43.2 kDa and 28.5 kDa in size, respectively. The mass of the

mutated variants of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, the MRSA *N*-acetylglucosamine-6-phosphate deacetylase and the MRSA glucosamine-6-phosphate deaminase was also verified by electrospray ionisation mass spectrometry.

3.2.3 The solution structure of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase

3.2.3.1. Quaternary structure as determined by analytical ultracentrifugation

Sedimentation velocity experiments using analytical ultracentrifugation demonstrate that wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase is dimeric in solution. The continuous sedimentation coefficient distribution [c(*s*)] distribution resulted in a single symmetrical peak, confirming that wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase is a single species in solution, with a sedimentation coefficient of 3.5 S (Figure 3.5 A). This data was then fitted to a continuous mass distribution [c(*M*)] model to calculate the apparent molecular mass of the species (Figure 3.5 B). The apparent molecular mass was determined to be 49.2 kDa. Considering the theoretical monomeric mass of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase is 24.5 kDa, this is consistent with the theoretical dimeric mass of 49 kDa. The absorbance at 285 nm versus radial position of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase shows a single sedimenting boundary, consistent with the presence of a single species (Figure 3.5 C). The randomly distributed residuals and low root-mean-square deviation (rmsd) (0.0031) indicate a good fit (Lebowitz *et al.*, 2002).

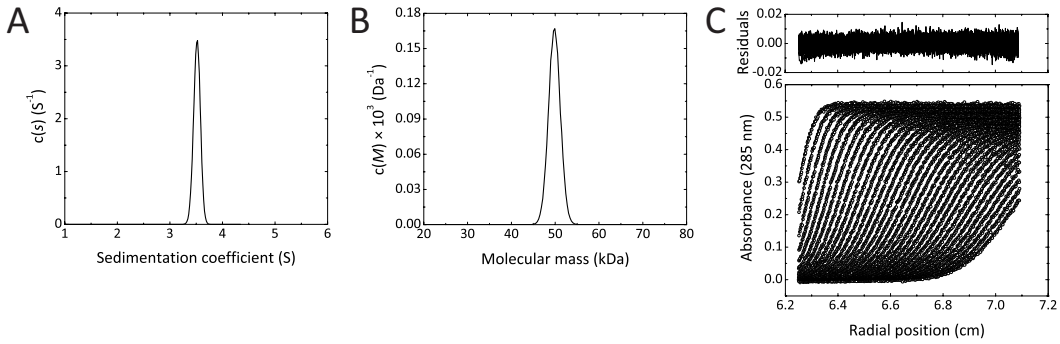


Figure 3.5. Sedimentation velocity analysis of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase using analytical ultracentrifugation. The experiment was conducted at 0.5 mg/mL (20.4 µM) of freshly purified protein in 20 mM 2-amino-2-hydroxymethyl-propane-1,3-diol (Tris), pH 8.0, 150 mM sodium chloride. Data were collected at 285 nm, 50 000 rpm and 20 °C. SEDNTERP was used to calculate the partial-specific volume of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase (0.7412 g/ml), solvent density (1.0029 g/ml) and viscosity (0.0102 poise) (Laue *et al.*, 1992). Data were fitted to a c(*s*) and c(*M*) model at a resolution of 300 and a confidence level of 0.95 using SEDFIT (Schuck, 2000). **(A)** The c(*s*) model plotted as a function of sedimentation coefficient. Data were fitted with an *s* value ranging between 1 and 6 S. The fit resulted in a frictional ratio (f/f_0) of 1.2928. **(B)** The c(*M*) model plotted as a function of molecular mass (kDa). Data were fitted with a minimum mass of 20 kDa and a maximum mass of 80 kDa. **(C)** Absorbance at 285 nm plotted as a function of radial position (cm). The raw data are represented as open symbols (o) and overlaid with the non-linear least squares best fit. The residuals for this fit are shown above.

3.2.3.2 Small angle X-ray scattering analysis of the solution structure

Small angle X-ray scattering (SAXS) data further verifies the dimeric solution structure of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. Guinier analysis was used to assess sample quality and the linearity of the plot confirms that the scattering is reliable (Figure 3.6 A). Scattering data are presented as an intensity plot (Figure 3.6 B) and a real space distance distribution function *P*(*r*) plot (Figure 3.6 C). The maximum dimension of the scattering particle (D_{max}) was calculated by *AUTOPOROD* (Petoukhov & Svergun, 2007) as 75.8 Å. The radius of gyration (R_g) determined by *GNOM* (Svergun, 1992) was calculated as 24.9 Å, in close agreement with the 23.5 Å R_g determined by Guinier analysis. The molecular mass extracted from the SAXS data is 40.4 kDa, which is slightly smaller than the 49 kDa theoretical dimeric mass and the 49.2 kDa apparent molecular mass derived from sedimentation velocity analysis. The low resolution of SAXS data would account for this discrepancy in size (Putnam *et al.*, 2007). The *P*(*r*) distribution

shows a single peak with a positively skewed tail at long distances, indicative that MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase acquires an elongated shape in solution.

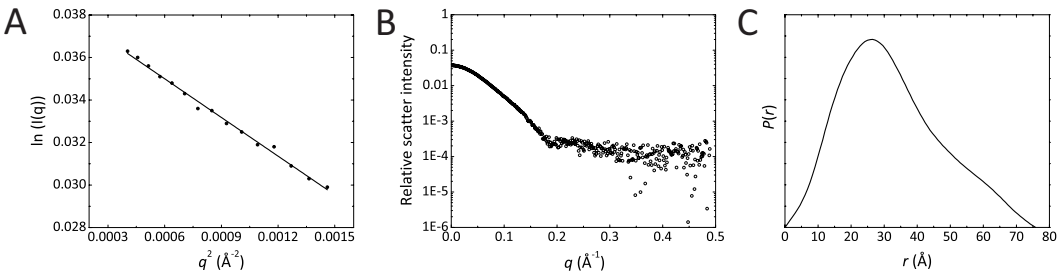


Figure 3.6. SAXS data for MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. (A) Guinier plot which is linear, determined using *GNOM*. **(B)** Experimental scattering data presented as an intensity plot. **(C)** *P*(*r*) plot determined using *GNOM*.

To determine whether the structure in solution accurately reflects my crystal structure (Section 3.2.8), the experimental scattering data were compared with the theoretical scattering generated from the crystal structure using CRYSOL (Svergun *et al.*, 1995). The theoretical scattering calculated for the crystal structure of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase fits the experimental scattering data ($\chi^2 = 0.56$), confirming that the crystal structure is an accurate representation of the solution structure (Figure 3.7).

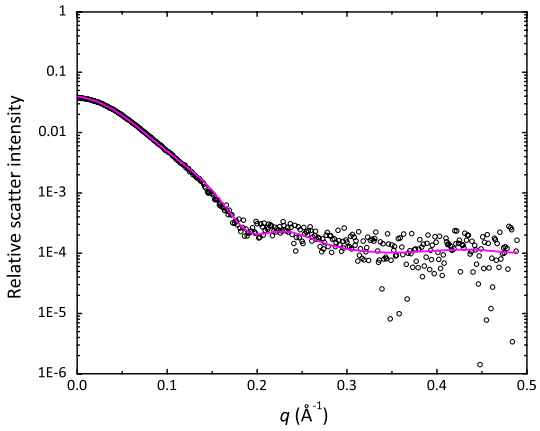


Figure 3.7. Comparison of experimental and theoretical scattering data. The experimental scattering profile of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase (presented as open circles, o) overlaid with the theoretical scattering profile of the crystal structure in magenta.

3.2.4 Development of a multi-enzyme kinetic assay for *N*-acetylmannosamine-6-phosphate 2-epimerase

As described in Section 3.1.4, an efficient method to characterise the activity of *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes is lacking. Thus, I developed a multi-enzyme coupled assay to determine the kinetic properties of *N*-acetylmannosamine-6-phosphate 2-epimerase in real time (Figure 3.8). This assay utilises the subsequent enzymes of the sialic acid degradation pathway, *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase, as well as phosphoglucoisomerase and glucose-6-phosphate dehydrogenase as coupling enzymes (Chapter Six, Section 6.9.2).

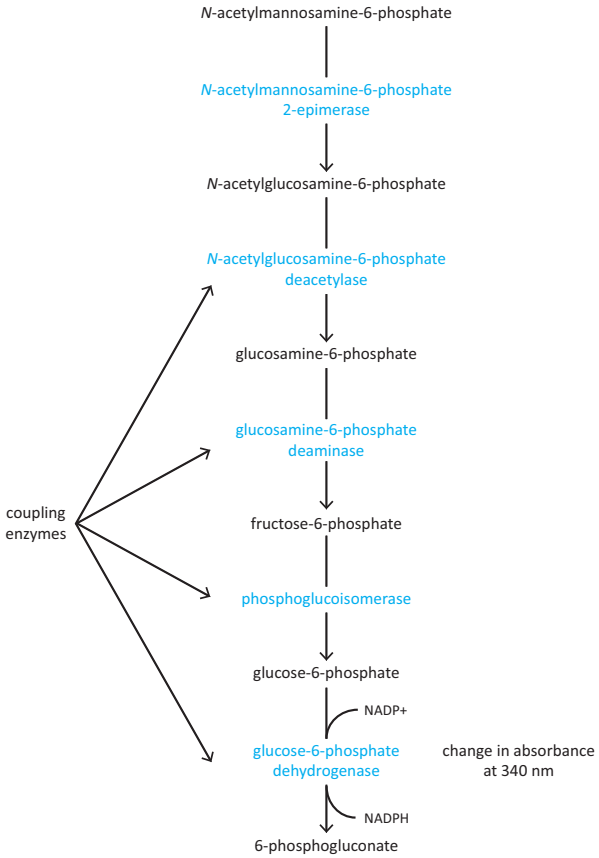


Figure 3.8. The multi-enzyme coupled assay developed for *N*-acetylmannosamine-6-phosphate 2-epimerase. As NADP+ is reduced into NADPH, the increase in absorbance at 340 nm is followed spectrophotometrically. The enzymes are labelled in blue.

Briefly, the assay is initiated by the addition of substrate (*N*-acetylmannosamine-6-phosphate), which is converted into *N*-acetylglucosamine-6-phosphate by *N*-acetylmannosamine-6-phosphate 2-epimerase. *N*-Acetylglucosamine-6-phosphate deacetylase, glucosamine-6-phosphate deaminase, phosphoglucoisomerase and glucose-6-phosphate dehydrogenase convert *N*-acetylglucosamine-6-phosphate into glucosamine-6-phosphate, fructose-6-phosphate, glucose-6-phosphate and 6-phosphogluconate, respectively. As glucose-6-phosphate dehydrogenase turns over glucose-6-phosphate into 6-phosphogluconate, nicotinamide adenine dinucleotide phosphate (NADP+) is simultaneously reduced to NADPH. Because NADP+ does not absorb at 340 nm and NADPH does, the increase in absorbance at this wavelength can be followed as the reaction is catalysed. The rate of this reaction will be proportional to the rate of catalysis by *N*-acetylmannosamine-6-phosphate 2-epimerase.

Prior to kinetic analysis, the secondary structure of the wild type and mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes was assessed to ensure that they were folded following purification (Section 3.2.4.1). In addition, the thermal stability of the wild type and the mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes, as well as the purified MRSA *N*-acetylglucosamine-6-phosphate deacetylase and MRSA glucosamine-6-phosphate deaminase coupling enzymes were evaluated (Section 3.2.4.2). Understanding the thermal stability of these enzymes was necessary to determine a suitable temperature by which this multi-enzyme coupled assay could be conducted.

Because this is a newly developed assay for kinetic analysis, various controls also had to be performed to ensure that the conditions of the experiment allowed for accurate and reproducible calculations to be made. Firstly, because multiple coupling enzymes were involved, it was important to ensure that none of these enzymes non-specifically acted upon the substrate, *N*-acetylmannosamine-6-phosphate (Section 3.2.4.3). In addition, it is important that all of the coupling enzymes were in excess so that they are not limiting the rate of the reaction being measured and the rate being measured is proportional to enzyme concentration (Section 3.2.4.4) (Cornish-Bowden, 2004).

3.2.4.1 Secondary structure of wild type and mutant MRSA N-acetylmannosamine-6-phosphate 2-epimerase enzymes

Circular dichroism (CD) spectroscopy was performed on the wild type and mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes to ensure that they were folded. In addition, this allowed for any differences in secondary structure between the enzymes to be assessed. The spectra recorded showed that the enzymes are folded (Figure 3.9). A broad minima is apparent between approximately 208 nm and 222 nm, which is characteristic of enzymes composed of both α-helices and β-strands.

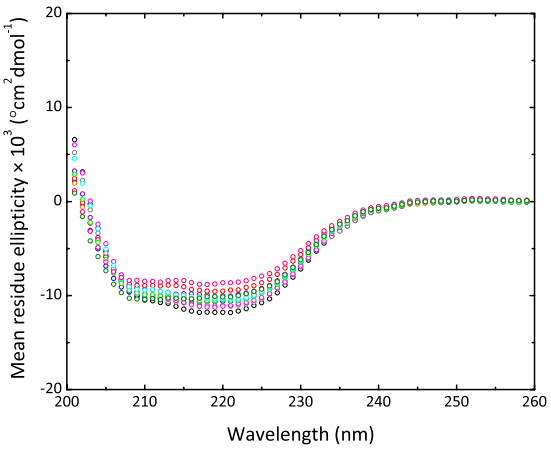


Figure 3.9. CD spectra of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes. CD spectroscopy was conducted at 0.1 mg/mL in 20 mM Tris, pH 8.0 for the wild type and mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes. Mean residue ellipticity is plotted as a function of wavelength. Data are represented as open circles, with wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase in black (o), Gln11Ala in blue (o), Gln11Ser in orange (o), Glu16Ala in purple (o), Glu16His in red (o), Arg40Ala in grey (o), Lys63Ala in magenta (o), Lys63Glu in lime green (o), Asp124Ala in cyan (o), Asp124Gln in pink (o) and Arg208Ala in dark green (o).

Although the spectra recorded follow a similar trend, they do not overlay with one another. Thus, small differences in secondary structure are apparent between the wild type and mutant enzymes. The percentage of each secondary structure component present in each enzyme is shown in Table 3.1. The wild type enzyme appears to have a slightly higher percentage of α-helices and a slightly lower percentage of β-strands/turns compared to the mutant enzymes. The mutated variants contain single amino acid mutations within the

active site of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. Although the exact location that the secondary structure is altered cannot be determined using CD spectroscopy, it is likely that these mutations are perturbing secondary structural elements at or near the active site of the enzyme. Crystal structures of the mutated variants would allow for changes in secondary structure to be probed, but this is beyond the scope of this thesis.

Table 3.1. The percentage of each secondary structure component in the wild type and mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes. Data was analysed using *CDNN* (Bohm *et al.*, 1992).

MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase	α-Helix (%)	β-Strand (%)	β-Turn (%)	Random coil (%)
Wild type	37	15	16	32
Gln11Ala	33	17	17	33
Gln11Ser	32	18	17	33
Glu16Ala	32	18	17	33
Glu16His	30	19	17	34
Arg40Ala	35	17	16	32
Lys63Ala	35	17	16	32
Lys63Glu	33	17	17	33
Asp124Ala	33	18	17	32
Asp124Gln	28	21	18	33
Arg208Ala	33	18	17	32

3.2.4.2 Thermal stability of the enzymes involved in the multi-enzyme coupled assay

The thermal stability and structural integrity of the wild type and mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes, as well as the purified MRSA *N*-acetylglucosamine-6-phosphate deacetylase and MRSA glucosamine-6-phosphate deaminase coupling enzymes was determined using differential scanning fluorimetry (Table 3.2). It is clear that the mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes have very similar melting temperatures to the wild type enzyme, suggesting that they are similarly folded into stable conformations. Moreover, the MRSA *N*-acetylglucosamine-6-phosphate deacetylase and the MRSA glucosamine-6-phosphate deaminase enzymes display a melting temperature of 47 °C and 43 °C, respectively.

For the commercially purchased phosphoglucoisomerase from *Saccharomyces cerevisiae* and glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides*, the literature

was consulted to ensure an appropriate temperature was selected for the assay that is also suitable for these coupling enzymes. Although the thermal stability has not been specifically measured, previous assays have shown that *S. cerevisiae* phosphoglucoisomerase is stable up to 60 °C (Maitra, 1971), whereas *L. mesenteroides* glucose-6-phosphate dehydrogenase is shown to be stable in assays conducted at room temperature (Vincent *et al.*, 2005; Zhu *et al.*, 2011). Taking this into consideration with the thermal stability of the wild type and mutated variants of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, the MRSA *N*-acetylglucosamine-6-phosphate deacetylase and the MRSA glucosamine-6-phosphate deaminase enzymes, a temperature of 25 °C was selected for kinetic analysis of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase.

Table 3.2. Melting temperature of the wild type and mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes, MRSA *N*-acetylglucosamine-6-phosphate deacetylase and MRSA glucosamine-6-phosphate deaminase. Differential scanning fluorimetry was conducted at 1 mg/mL in 20 mM Tris, pH 8.0.

Enzyme	Melting temperature (°C)
MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase	
Wild type	40.0
Gln11Ala	40.0
Gln11Ser	39.3 ± 0.3
Glu16Ala	40.7 ± 0.3
Glu16His	39.0
Arg40Ala	41.3 ± 0.3
Lys63Ala	43.0
Lys63Glu	40.7 ± 0.3
Asp124Ala	44.0
Asp124Gln	38.7 ± 0.3
Arg208Ala	42.0
MRSA *N*-acetylglucosamine-6-phosphate deacetylase	47.0
MRSA glucosamine-6-phosphate deaminase	43.0

3.2.4.3 Ensuring that the coupling enzymes do not catalyse their reaction with N-*acetylmannosamine-6-phosphate*

A test was conducted to ensure that the coupling enzymes (MRSA *N*-acetylglucosamine-6-phosphate deacetylase, MRSA glucosamine-6-phosphate deaminase, *S. cerevisiae* phosphoglucoisomerase and *L. mesenteroides* glucose-6-phosphate dehydrogenase) do not non-specifically catalyse their reaction with *N*-acetylmannosamine-6-phosphate, the substrate for *N*-acetylmannosamine-6-phosphate 2-epimerase. To do this, the baseline

absorbance of the assay mixture (100 mM Tris, pH 8.0, 5 mM magnesium chloride, 1 mM NADP+) with substrate, *N*-acetylmannosamine-6-phosphate, was recorded without any of the coupling enzymes (Figure 3.10). The coupling enzymes were sequentially added to the mixture in reverse order (glucose-6-phosphate dehydrogenase, phosphoglucoisomerase, glucosamine-6-phosphate deaminase and *N*-acetylglucosamine-6-phosphate deacetylase, respectively) to ensure that they were not turning over the substrate.

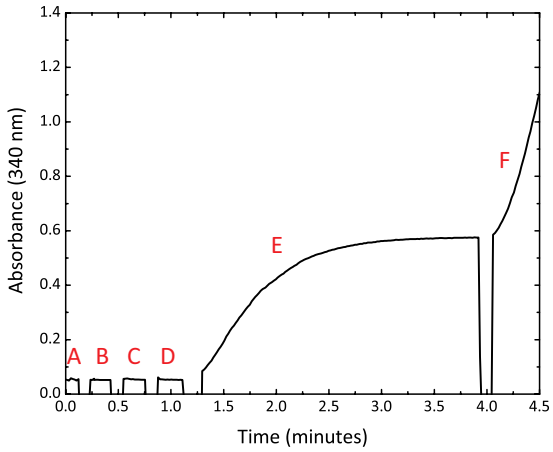


Figure 3.10 The absorbance at 340 nm as each enzyme is sequentially added to the assay mixture. The assay mixture contains 100 mM Tris, pH 8.0, 5 mM magnesium chloride, 1 mM NADP+ and 5 mM of substrate (*N*-acetylmannosamine-6-phosphate 2-epimerase) for *N*-acetylmannosamine-6-phosphate 2-epimerase. The baseline absorbance at 340 nm of the assay mixture does not increase (A). When the glucose-6-phosphate dehydrogenase, phosphoglucoisomerase and glucosamine-6-phosphate deaminase coupling enzymes are added to the assay (B, C and D, respectively), the absorbance at 340 nm also does not increase. When the *N*-acetylglucosamine-6-phosphate deacetylase coupling enzyme is added to the assay, there is an increase in absorbance at 340 nm, which eventually flattens. When *N*-acetylmannosamine-6-phosphate 2-epimerase (F) is added, the absorbance at 340 nm steeply increases.

My control showed that as the coupling enzymes were sequentially added to the assay, the baseline absorbance did not change until *N*-acetylglucosamine-6-phosphate deacetylase was added, which was the final coupling enzyme to be tested. This means that a change in absorbance at 340 nm becomes apparent as glucose-6-phosphate dehydrogenase converts NADP+ into NADPH. Thus, *N*-acetylglucosamine-6-phosphate deacetylase is either conducting its reaction with *N*-acetylmannosamine-6-phosphate, or it is utilising it's own substrate, *N*-acetylglucosamine-6-phosphate, which is contaminating the assay mixture.

Because the commercially purchased *N*-acetylmannosamine-6-phosphate is not completely pure (minimum purity 95%), there may be a small percentage of its epimer, *N*-acetylglucosamine-6-phosphate present when made up in solution. Importantly, *N*-acetylglucosamine-6-phosphate is the substrate for *N*-acetyglucosamine-6-phosphate deacetylase. As shown in Figure 3.10, the increase in absorbance at 340 nm associated with the addition of *N*-acetylglucosamine-6-phosphate deacetylase eventually flattens, implying that there is effectively no substrate left in the assay mixture for this enzyme. Then, when the *N*-acetylmannosamine-6-phosphate 2-epimerase enzyme is added to the assay mixture, the absorbance at 340 nm again increases. Thus, the *N*-acetylglucosamine-6-phosphate deacetylase enzyme cannot be utilising *N*-acetylmannosamine-6-phosphate as a substrate and activity can be attributed to a small amount of *N*-acetylglucosamine-6-phosphate being present in the commercially purchased *N*-acetylmannosamine-6-phosphate.

The presence of *N*-acetylglucosamine-6-phosphate in the commercially purchased *N*-acetylmannosamine-6-phosphate means that if the amount of *N*-acetylmannosamine-6-phosphate added to the assay is doubled, the amount of *N*-acetylglucosamine-6-phosphate also doubles. This combined with the absorbance range of the spectrophotometer, which is linear between approximately 0 and 1.2, defined an upper limit for the amount of *N*-acetylmannosamine-6-phosphate that could be used. This upper limit was determined as 5 mM *N*-acetylmannosamine-6-phosphate, by which kinetic analysis of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes could not be conducted above.

3.2.4.4 Ensuring that the coupling enzymes are in excess and the rate of the reaction is proportional to MRSA N-acetylmannosamine-6-phosphate 2-epimerase concentration

As described in Chapter Two, Section 2.2.3.1, it is important that all of the coupling enzymes are in excess so that they are not limiting the rate of the reaction being measured. In addition, the initial rate of the reaction that is being measured must be proportional to enzyme concentration (Cornish-Bowden, 2004). To test that the enzymes are in excess, the amount of each coupling enzyme initially added to the assay was sequentially doubled in reverse order (glucose-6-phosphate dehydrogenase, phosphoglucoisomerase, glucosamine-6-phosphate deaminase and *N*-acetylglucosamine-6-phosphate deacetylase, respectively) to ensure that the rate of the reaction did not change. This control determined that 50 μg of *N*-

acetylglucosamine-6-phosphate deacetylase, 100 µg of glucosamine-6-phosphate deaminase, 50 µg of phosphoglucoisomerase and 25 µg of glucose-6-phosphate dehydrogenase per assay could be used as excess. At these concentrations, the rate of the reaction did not change when the concentration of each coupling enzyme was doubled (Figure 3.11 A). To ensure that the initial rate of the reaction is proportional to the concentration of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, a range of concentrations were analysed between 0.25 µg and 2 µg per assay. A final concentration of 0.25 µg was selected for kinetic analysis of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, as this concentration was proportional to the initial rate of the reaction, depicted by the linearity of the initial velocity observed in Figure 3.11 B.

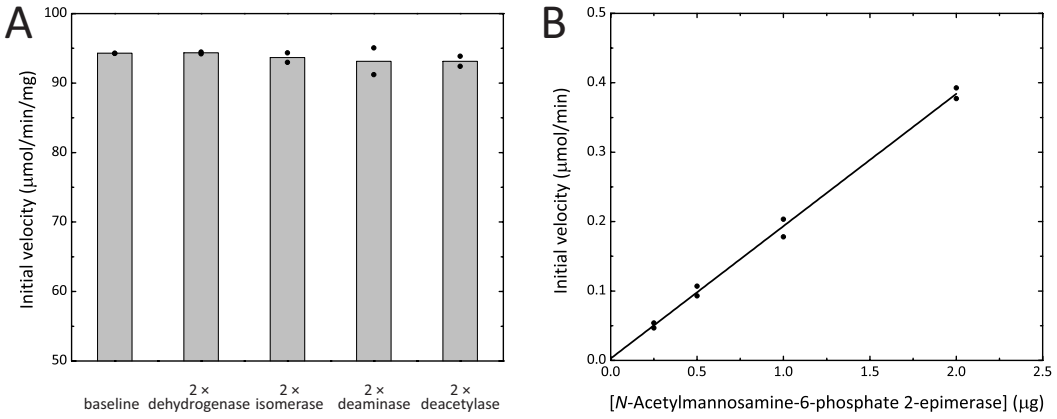


Figure 3.11. Excess and concentration control for the multi-enzyme coupled assay developed for MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. (A) The initial velocity of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase (baseline) with 50 µg of *N*-acetylglucosamine-6-phosphate deacetylase, 100 µg of glucosamine-6-phosphate deaminase, 50 µg of phosphoglucoisomerase and 25 µg of glucose-6-phosphate dehydrogenase. The amount of each enzyme was sequentially doubled in the order depicted in the graph. The mean is shown as a bar and the data is plotted as data points. **(B)** The initial velocity of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase as a function of enzyme concentration. Data were fitted to a linear model resulting in an R^2 value of 0.99.

3.2.5 Michaelis-Menten kinetic analysis of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase

The initial rate of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase with varying concentrations of *N*-acetylmannosamine-6-phosphate was measured and fitted to the Michaelis-Menten kinetic model, displaying a hyperbolic dependence of initial velocity on substrate concentration (Figure 3.12). For the wild type enzyme, the Michaelis-Menten constant (K_M) value was found to be 0.7 ± 0.03 mM, the maximum velocity (V_{max}) was calculated as 102 ± 1 μmol/min/mg and the catalytic turnover number (k_{cat}) was determined to be 41.6 ± 0.6 s^{-1}. ^{1}H NMR kinetic analysis of *N*-acetylmannosamine-6-phosphate 2-epimerase from *C. perfringens* revealed a K_M value of 4.5 mM (Pelissier *et al.*, 2014). Thus, the wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzyme has a much higher affinity for *N*-acetylmannosamine-6-phosphate than that of *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase.

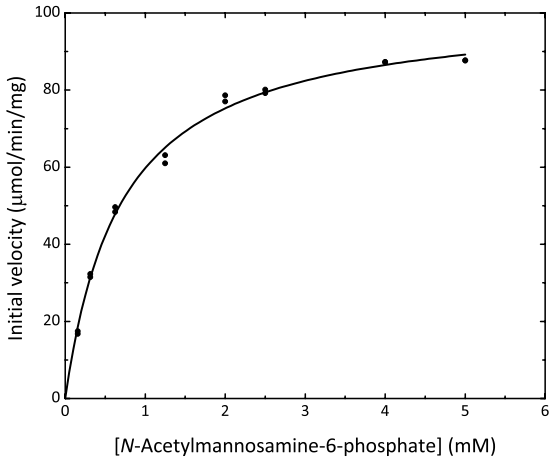


Figure 3.12. Michaelis-Menten kinetic analysis of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. Data were fitted to the Michaelis-Menten equation with an R^2 value of 0.99.

3.2.6 Crystallisation of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase

To further investigate the solution structure of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, crystallisation studies were conducted (Chapter 6, Section 6.7.1). The JCSG+ Suite produced crystals in several of the screen conditions. Condition E1 (1 M trisodium citrate, 0.1 M sodium cacodylate, pH 6.5) produced a cluster of crystals at 8 °C. These crystals were suitable for the collection of high resolution X-ray diffraction data to 1.84 Å, without any optimisation required.

3.2.7 Data processing and structure refinement of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase

The data was processed in the space group $P2_12_12$ and solved by molecular replacement using the monomer of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase (PDB entry 1y0e) as the search model. The presence of two screw axes was verified by the presence of systematic absences in only two of the axes. Consistent with the solution structure, the asymmetric unit contains two monomers that correspond to a dimer. My construct differs by two amino acid residues to that of PDB entry 1y0e. Firstly, PDB entry 1y0e contains an extra alanine residue at position one, which precedes the methionine located at position one of the construct presented. Secondly, a valine residue located at position 220 of PDB entry 1y0e corresponds to an isoleucine residue located at position 219 of the construct presented. Following multiple rounds of structure refinement, the residual factor (R_{factor}) and the free R_{factor} (R_{free}) were reduced to 16.9 and 19.8%, respectively. This model includes all residues except for the methionine at position one of each monomer, which was too disordered to model into the density at a sigma level of one. All relevant data collection and processing parameters as well as the structure and refinement statistics are provided in Table 3.3.

Table 3.3. X-ray data collection and refinement statistics for wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. Statistical values for the highest resolution shells are given in parentheses. Each monomer within the asymmetric unit has a molecular weight of 49 090 Da.

Statistics	MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase
data collection statistics	
wavelength (Å)	0.95369
number of images	360
oscillations (°)	0.5
space group	$P2_12_12$
unit cell parameters (Å, °)	$a = 77.1$, $b = 173.7$, $c = 44.4$, $\alpha = \beta = \gamma = 90$
resolution range (Å)	46.3-1.84 (1.88-1.84)
observed reflections	189024 (9200)
unique reflections	52704 (3125)
mean $I/\sigma(I)$	10.8 (1.8)
completeness (%)	99.8 (98.5)
R_{merge}†	0.082 (0.549)
$R_{r.i.m.}$‡	0.112 (0.738)
$R_{p.i.m.}$§	0.076 (0.490)
wilson B value (Å^2)	9.6
molecules per asymmetric unit	2
V_M (Å^3 Da^{-1})	3.04
solvent content (%)	59.5
structure and refinement statistics	
R_{factor} (%)	16.9
R_{free} (%)	19.8
number of atoms	
protein	3422
water	531
ligands	18
rmsd	
bonds (Å)	0.015
angles (°)	1.524
average B factors (Å^2)	
protein	15.9
water	26.6
ligands	32.4
Ramachandran plot, residues (%)	
favoured region	96.8
allowed region	2.74
disallowed region	0.46

R_{merge} † = $\sum_{hkl} \sum_i |I_i\,(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i I_i\,(hkl)$.
$R_{r.i.m}$ ‡ = $\sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i\,(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i I_i\,(hkl)$.
$R_{p.i.m.}$§ = $\sum_{hkl} [1/(N-1)]^{1/2} \sum_i |I_i\,(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i I_i\,(hkl)$.

3.2.8 The structure of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase

Prior to the beginning of this thesis, crystal structures of *N*-acetylmannosamine-6-phosphate 2-epimerase from MRSA (PDB entry 1y0e; Midwest Centre for Structural Genomics, unpublished work), *Salmonella enterica* (PDB entry 3igs; Centre for Structural Genomics of Infectious Diseases, unpublished work) and *Streptococcus pyogenes* (PDB entry 1yxy; Midwest Centre for Structural Genomics, unpublished work) had been solved and deposited in the PDB, but not yet published in the literature. Although the structure had already been solved for MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, a crystallisation protocol was developed for my construct in particular, which allowed for the structure to be solved. Having this protocol and a ligand free crystal structure available is important for studies directed at understanding how the chemistry is catalysed by these enzymes as well as for the development of inhibitors in the future.

Consistent with the solution based structure, the crystal structure of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase is a dimer, consisting of two single domain triosephosphate isomerase barrels [$(\beta/\alpha)_8$-barrels] (Figure 3.14 A). Unlike any other known $(\beta/\alpha)_8$-barrel enzymes, dimerisation between monomers is enabled by exchanging the C-terminal α-helix, which projects away from each barrel and packs against the β-sheet of the neighbouring monomer. This arrangement is highly similar to that observed for all structures solved to date of *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes. The monomer of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase is shown in Figure 3.14 B, with the consecutive β-strands labelled *a* through *h* and the α-helices labelled *A* through *H*. It is α-helix *H* that is exchanged to enable dimerisation between the monomers. As with all $(\beta/\alpha)_8$-barrel enzymes, the active site is located in the centre of the barrel at the C-terminal end of the β-strands (Wierenga, 2001).

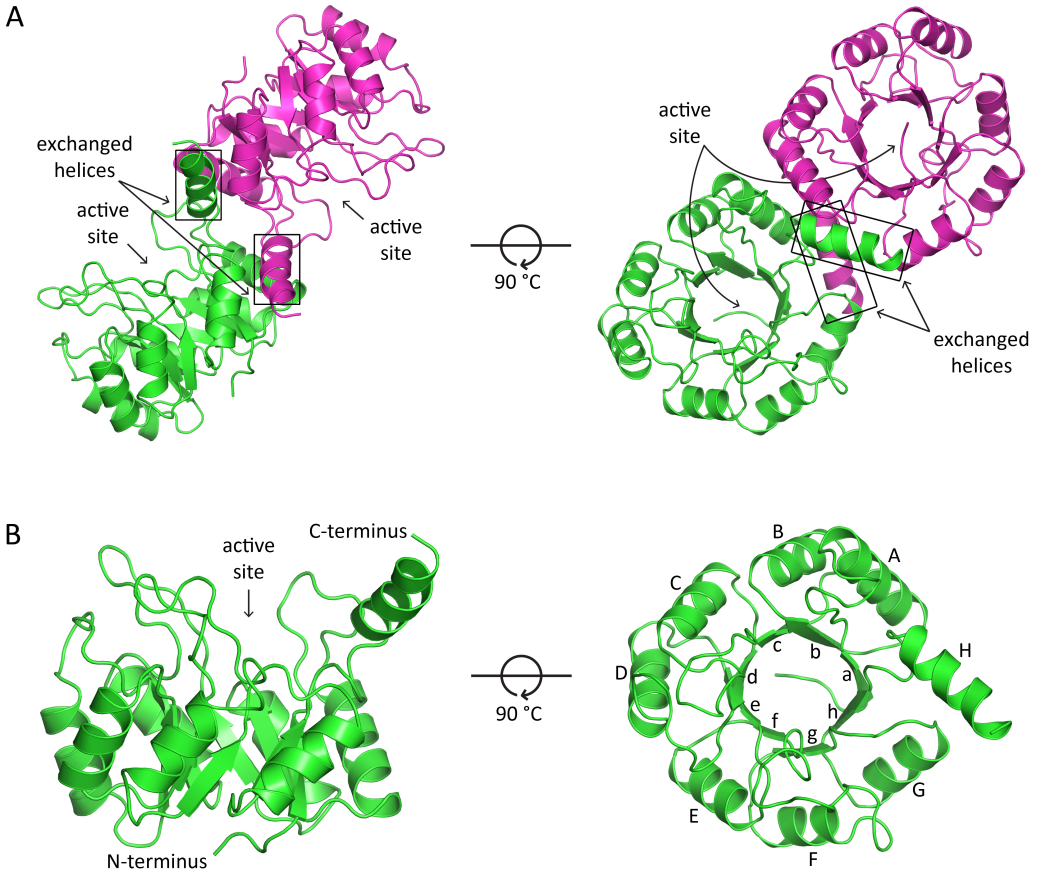


Figure 3.14. The crystal structure of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. (A) The dimeric structure of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase formed by exchanging an α-helix between monomers. **(B)** The monomeric structure of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase with the consecutive β-strands labelled *a* through *h* and the α-helices labelled *A* through *H*.

Sequence analysis of *N*-acetylmannosamine-6-phoshate enzymes demonstrates a number of strictly conserved residues among gram positive bacterial species (Figure 3.15). Based on this conservation and considering the position of active sites within $(\beta/\alpha)_8$-barrel enzymes (Wierenga, 2001), my structure of MRSA *N*-acetylmannosamine-6-phosphate could be used to identify residues that may be important for catalysis (Figure 3.16). Specifically, residues Gln11, Glu16, Arg40, Lys63, Asp124 and Arg208 may have a role in catalysis, since they line the cleft at the C-terminal ends of the β-strands that is proposed to form the active site. To elucidate whether these residues are important for the chemistry catalysed by MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, they were mutated to an alanine, producing Gln11Ala, Glu16Ala, Arg40Ala, Lys63Ala, Asp124Ala and

Arg208Ala. In addition, Gln11 was also mutated into a serine, producing Gln11Ser; this is a more conservative amino acid substitution, changing the length of the side chain but retaining the ability to form hydrogen bonds. Glu16 was also mutated into a histidine, producing Glu16His, changing the charge of the amino acid side chain from negative to positive. Similarly, Lys63 was also mutated into a glutamate, producing Lys63Glu, also changing the charge of the amino acid side chain, but this time from positive to negative. Asp124 was mutated into a glutamine, producing Asp124Gln and changing the amino acid side chain from a negative charge to an uncharged functionality. As shown in Figure 3.16, Arg40 forms a salt bridge with Asp124 and removing the charge at this position will disturb the ability of these residues to form a salt bridge.

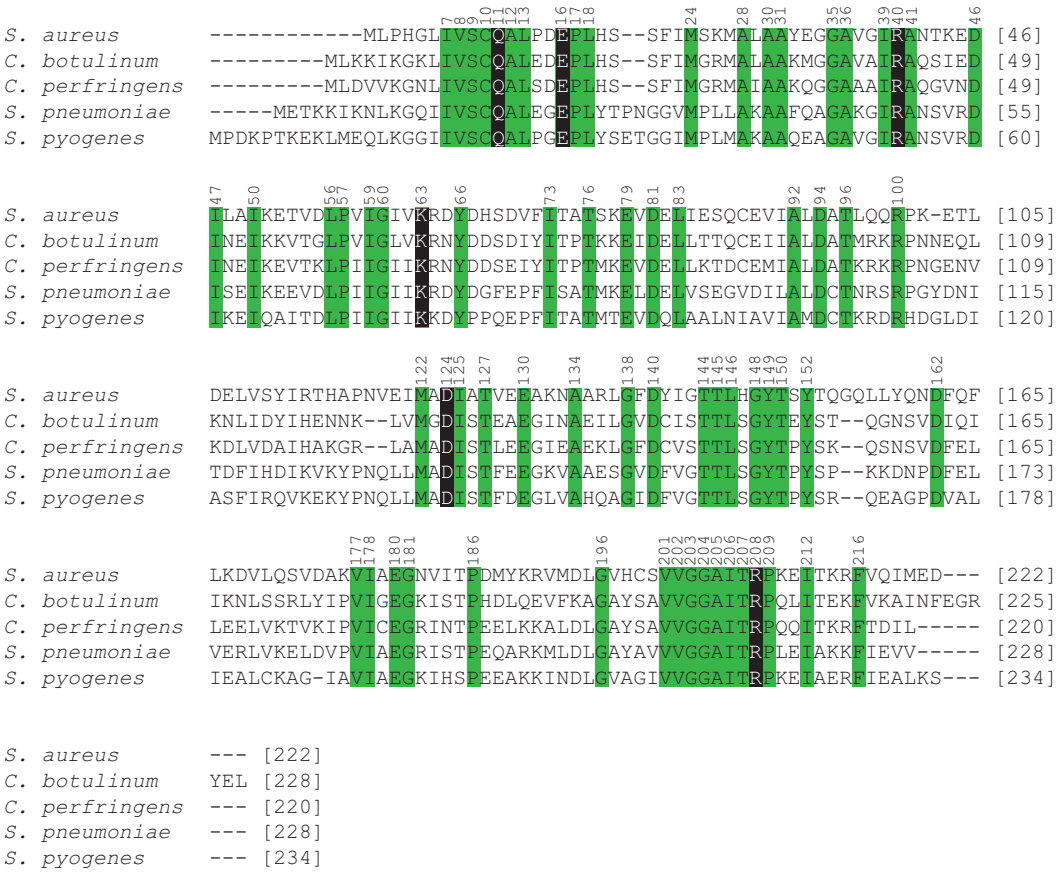


Figure 3.15. Sequence alignment of *N*-acetylmannosamine-6-phosphate 2-epimerase from five gram positive species of bacteria. Conserved residues are highlighted in green, while those that may be involved in catalysis are highlighted in black. Species include *S. aureus*, *Clostridium botulinum*, *C. perfringens*, *Streptococcus pneumoniae* and *S. pyogenes*. Conserved residues are numbered according to *S. aureus*. The alignment was generated using *ClustalW* (Larkin *et al.*, 2007) (http:// www.genome.jp/tools/clustalw/).

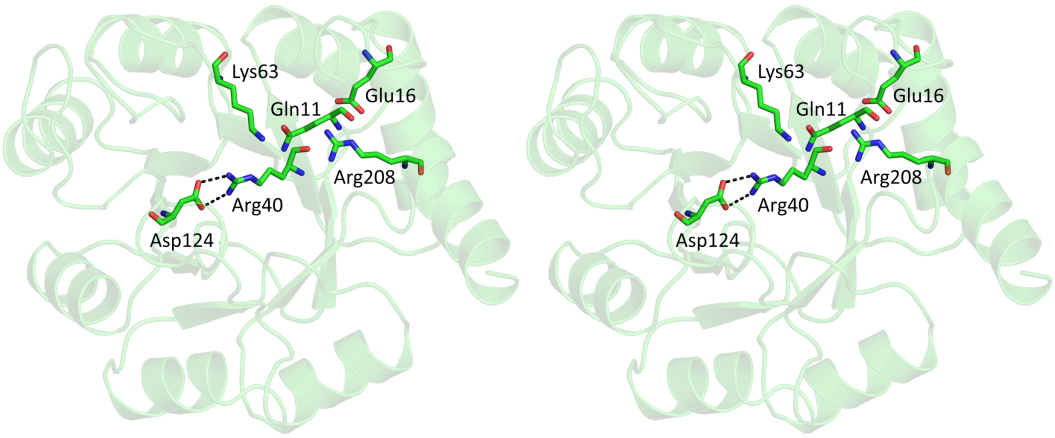


Figure 3.16. Potential active site residues of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. A stereo view of the MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase monomer. The positions of the potential catalytic residues, Gln11, Glu16, Arg40, Lys63, Asp124 and Arg208 are shown. The bonds forming the salt bridge between Arg40 and Asp124 are indicated by black dashes.

3.2.9 Michaelis-Menten kinetic analysis of mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes

Because enzyme kinetics in combination with mutagenesis can be used to test hypothetical catalytic mechanisms (Cornish-Bowden, 2004), kinetic analysis was used to assess whether the residues identified in Section 3.2.8 are important for catalysis in MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. For each of the mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes, initial rate data was collected using the multi-enzyme coupled assay that I developed and then fitted to the Michaelis-Menten kinetic model. All kinetic parameters for the wild type and mutated variants of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase can be found in Table 3.4. Because the kinetic assay could not be performed above 5 mM of *N*-acetylmannosamine-6-phosphate, the kinetic parameters for the mutant *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes that displayed some activity carry significant error. However, it is likely that the catalytic activity of these mutants is exceptionally low compared to the wild type and the K_M parameters are significantly higher than 5 mM.

Table 3.4. Kinetic parameters for wild type and mutated variants of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. The V_{max}, k_{cat} and K_M calculated in OriginPro.

MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase	V_{max} (μmol/min/mg)	k_{cat} (s^{-1})	Relative k_{cat} (%)	K_M (mM)
Wild type	101.7 ± 1.4	41.6 ± 0.6	100	0.7 ± 0.03
Gln11Ala	11.5 ± 2.1	4.7 ± 0.9	11.3	10.3 ± 2.6
Gln11Ser	33.1 ± 20.1	13.5 ± 8.2	32.5	50.0 ± 32.9
Glu16Ala	2.9 ± 0.1	1.2 ± 0.1	2.9	20.1 ± 0.7
Glu16His	10.7 ± 9.6	4.4 ± 3.9	10.6	93 ± 87
Arg40Ala	NA	-	-	-
Lys63Ala	NA	-	-	-
Lys63Glu	NA	-	-	-
Asp124Ala	NA	-	-	-
Asp124Gln	NA	-	-	-
Arg208Ala	3.4 ± 1.8	1.4 ± 0.7	3.4	31.6 ± 19.0

*NA denotes no activity

The Arg40Ala, Lys63Ala, Lys63Glu, Asp124Ala and Asp124Gln mutated variants showed no detectable activity, even at enzyme concentrations at least 10^3 times more than that of the wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase. The Gln11Ala, Gln11Ser, Glu16Ala, Glu16His and Arg208 mutated variants displayed catalytic efficiencies that were significantly lower than the wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase (Figure 3.17). Thus, Lys63 and the salt bridge formed between Arg40 and Asp124 are essential for the chemistry catalysed by MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, while Gln11, Glu16 and Arg208 also participate in catalysis. Participation may involve substrate binding and/or binding to the residues implicated in catalysis to stabilise the substrate and/or active site.

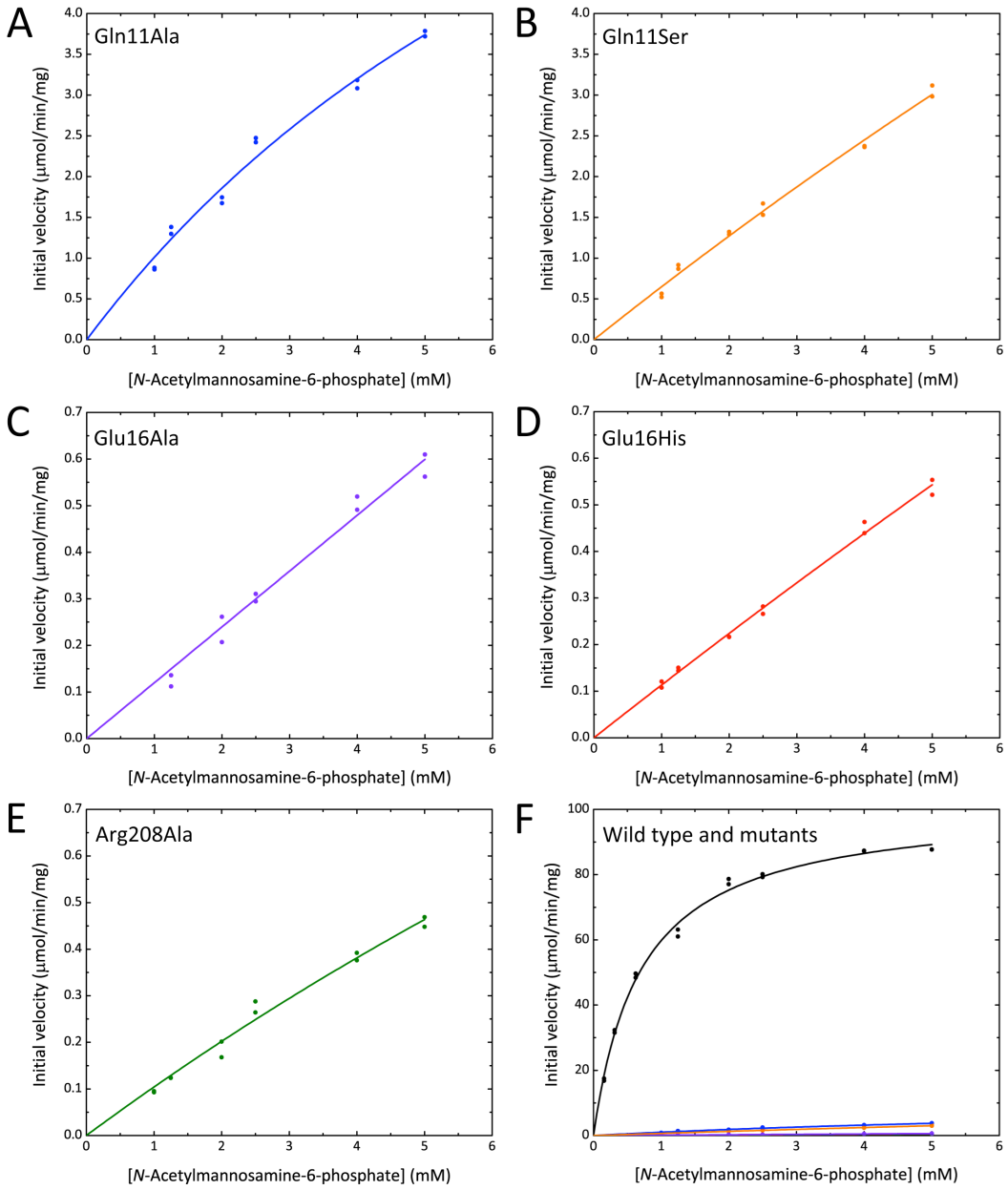


Figure 3.17. Michaelis-Menten kinetic analysis of the mutant MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes. Data were fitted to the Michaelis-Menten equation. **(A)** Gln11Ala had an R^2 value of 0.98. **(B)** Gln11Ser had an R^2 value of 0.99. **(C)** Glu16Ala had an R^2 value of 0.96. **(D)** Glu16His had an R^2 value of 0.99. **(E)** Arg208Ala had an R^2 value of 0.98. **(F)** The wild type (black) combined with the mutated variants.

3.2.10 The structure and proposed mechanism of *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase

A structural and ^{1}H NMR kinetic analysis of wild type and mutated variants of *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase was published during the course of this study (Pelissier *et al.*, 2014). As described in Section 3.1.3, this analysis identified residues that are essential for catalysis by *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase. Of the residues identified in Section 3.2.8, Lys66, Arg43 and Asp126 of *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase (equivalent to Lys63, Arg40 and Asp124 in MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase) were determined to be crucial for the chemistry catalysed by this enzyme. In addition, Glu180 (also equivalent to Glu180 in MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase) was shown to participate in catalysis, but was not essential, as a Glu180Ala mutant displayed 1.9% activity relative to the wild type.

The structures of *C. perfringens N*-acetylmanosamine-6-phosphate with bound product (*N*-acetylglucosamine-6-phosphate) and a Lys66Ala mutant with unprocessed substrate (*N*-acetylmannosamine-6-phosphate) demonstrate that Lys66 is ideally positioned relative to the C2 site of epimerisation. In addition, the guanidium group of Arg43, which forms a salt bridge with Asp126, is proposed to hydrogen bond with the O1 aldehyde of the substrate at position C1. As described in Section 3.1.3, Pelissier *et al.* (2014) propose that epimerisation occurs *via* a deprotonation/reprotonation mechanism, whereby the conserved lysine abstracts a proton from the C2 position of *N*-acetylmannosamine-6-phosphate and generates the epimer by reprotonation of the opposite face. To do this, the lysine is proposed to act as a Brønsted base on one stereoisomer and an acid catalyst on the other.

3.2.11 Alternative catalytic mechanisms

A critical analysis of the *C. pefringens N*-acetylmannosamine-6-phosphate 2-epimerase structures reveals that the proposed deprotonation/reprotonation is an unlikely catalytic mechanism. The active sites of my MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase structure and the *N*-acetylmannosamine-6-phosphate and *N*-acetylglucosamine-6-phosphate bound structures of *C. perfringens N*-acetylmannosamine-6-phosphate 2-

epimerase are overlaid in Figure 3.18. Interestingly, the side chain orientations of all catalytically important residues are highly similar in each structure. Thus, the active site of *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes is analogous between organisms and the presence of substrate or product does not affect the overall tertiary structure of the enzyme.

As depicted in Figure 3.18, Lys63 (numbered according to MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase) is in a position where it could abstract a proton from the C2 position of one stereoisomer, but it is not clear as to how it could reprotonate the opposite face to generate the epimer. For this to happen, Pelissier *et al.* (2014) suggest that following proton abstraction, the negatively charged intermediate must perform a 45 to 50° motion around the C2-C3 bond. In addition, the Gln11 side chain (equivalent to Gln14 in *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase), adopts two alternate conformations in the *C. perfringens N*-acetylglucosamine-6-phosphate bound structure. In one of these conformations, there is room for the lysine to move, which would accommodate the aldehyde group of the intermediate following rotation of the C2-C3 bond. If the lysine did move to this position, then Pelissier *et al.* (2014) propose that reprotonation on the opposite face could occur. However, no conformational rotamers of this lysine were identified in any structure, so this notion is contentious.

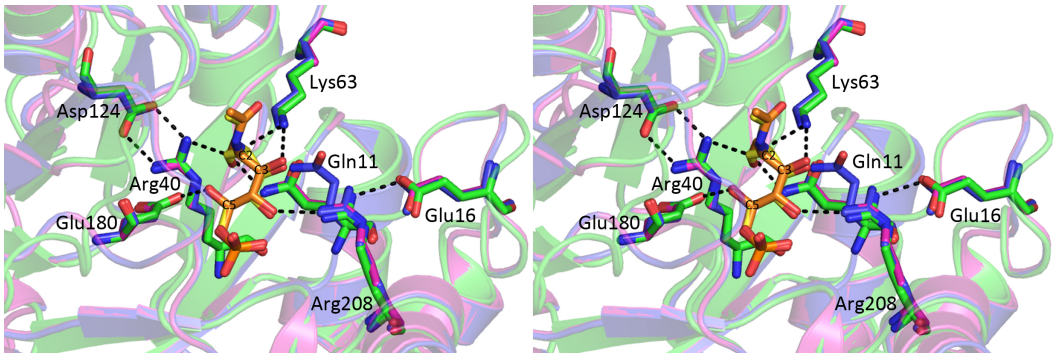


Figure 3.18. Alignment of the active sites. A stereo view of the MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase (green), *N*-acetylglucosamine-6-phosphate bound *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase (blue) and *N*-acetylmannosamine-6-phosphate bound *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase (magenta) active sites reveal a conserved architecture. The bonds that could be formed with *N*-acetylmannosamine-6-phosphate (yellow) and *N*-acetylglucosamine-6-phosphate (orange), as well as the salt bridge between Arg40 and Asp124 are indicated by black dashes.

Alternatively, I suggest that the catalytic mechanism of *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes may occur *via* the formation of a Schiff base or a proton displacement mechanism mediated by the substrate. For *C. perfringens* and MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes, the highly conserved and catalytically essential Lys63 residue is in a position where epimerisation *via* Schiff base formation is possible (Figure 3.19 A). This type of mechanism is already seen for *N*-acetylneuraminate lyase (Chapter Two) (Lawrence *et al.*, 1997). The presence of a Schiff base during catalysis by MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase can be tested with sodium borohydride, which will reduce the Schiff base and inactivate the enzyme (Section 3.2.12).

For a proton displacement mechanism mediated by the substrate (Figure 3.19 B), I propose that Lys63 deprotonates the C2 site and the substrate hydroxyl on the C5 position mediates reprotonation of the opposite face, which is ideally placed to do so and is activated by Glu180. Importantly, this does not require the substrate and the active site to reorient, which is consistent with the available crystallographic data. It is unclear whether the mechanism is concerted, or whether it involves an intermediate enolate form. But because the salt bridge formed between Arg40 and Asp124 is crucial for catalysis by both *C. perfringens* and MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, the formation of an enolate intermediate is highly likely. As shown in Figure 3.18, the guanidium of Arg40 can hydrogen bond with the O1 of the aldehyde at the C1 position, which would be negatively charged in an enolate intermediate and thus stabilised by Arg40. We note that the structure of the substrate and product in the active site is ideally orientated for this mechanism and the Lys63 residues and the substrate C5 hydroxyl are perfectly placed to either abstract or protonate the C2 carbon.

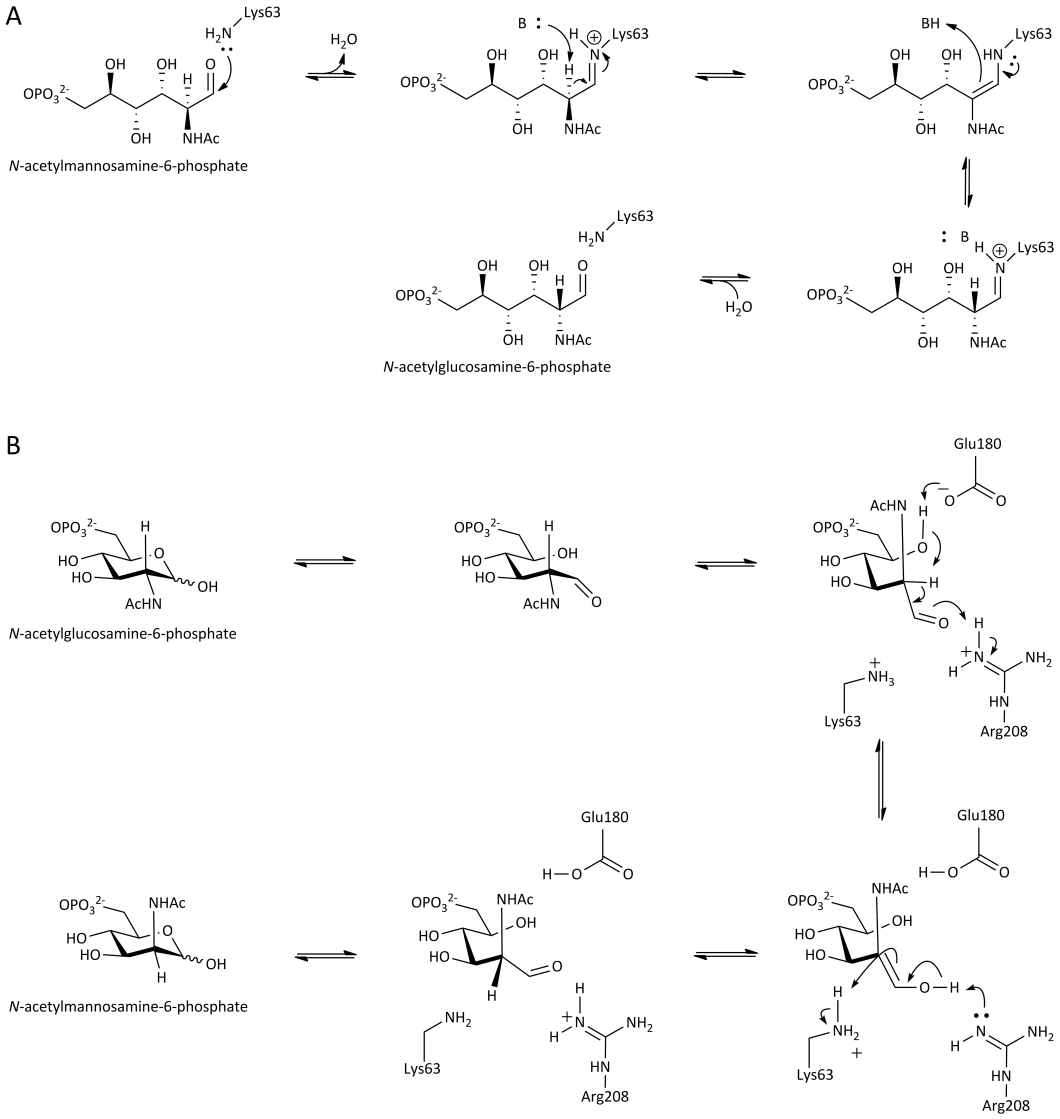


Figure 3.19. Alternative catalytic strategies for *N*-acetylmannosamine-6-phosphate 2-epimerase. (A) Epimerisation may occur *via* the formation of a Schiff base with Lys63. **(B)** Epimerisation may occur *via* a proton displacement mechanism mediated by the substrate.

3.2.12 Sodium borohydride reduction of wild type MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase

The formation of a Schiff base during catalysis can be observed by enzyme inactivation in the presence of sodium borohydride and substrate. For *N*-acetylneuraminate enzymes and dihydrodipicolinate synthase enzymes, Schiff base formation between a highly conserved

lysine residue and pyruvate is reduced by sodium borohydride, resulting in complete inactivation of the enzyme (Mirwaldt *et al.*, 1995; Lawrence *et al.*, 1997). To test the hypothesis that the mechanistic strategy employed by MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase may involve the formation of a Schiff base, the enzyme was incubated in the presence of sodium borohydride and its substrate, *N*-acetylmannosamine-6-phosphate (Chapter Six, Section 6.9.3). As a positive control, MRSA *N*-acetylneuraminate lyase (Chapter Two) was also incubated in the presence of sodium borohydride and its substrate, sialic acid. To determine whether MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase and MRSA *N*-acetylneuraminate lyase were inactivated in the presence of sodium borohydride, their respective coupled enzyme assays were used to assess activity. As shown in Table 3.5, MRSA *N*-acetylmannosamine-6-phosphate was not inactivated in the presence of sodium borohydride and substrate. But complete inactivation of MRSA *N*-acetylneuraminate lyase was seen in the presence of sodium borohydride and substrate. Thus, the mechanistic strategy of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase is unlikely to involve the formation of a Schiff base.

Table 3.5. Sodium borohydride reduction of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase and MRSA *N*-acetylneuraminate lyase. Enzymatic activity of these enzymes without sodium borohydride and following treatment with sodium borohydride in the presence of substrate.

Enzyme condition	Activity (%)
MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase	100
MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase + sodium borohydride + substrate	97
MRSA *N*-acetylneuraminate lyase	100
MRSA *N*-acetylneuraminate lyase + sodium borohydride + substrate	0

3.3 Summary

This chapter investigates the structure and catalytic mechanism of *N*-acetylmannosamine-6-phosphate 2-epimerase from MRSA. A new assay is developed that reports *N*-acetylmannosamine-6-phosphate 2-epimerase activity in real time, which will be convenient and suitable for testing inhibitors, unlike the current NMR assay that is reported (Pelissier *et al.*, 2014). The assay provides high quality rate data to determine kinetic constants for MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase, which suggest that this enzyme has a much higher affinity for its substrate compared to *C. perfringens N*-acetylmannosmaine-6-phosphate 2-epimerase.

Solution studies suggested that the enzyme is dimeric in its quaternary architecture, consistent with the crystal structure. Although there were some structures for *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes solved prior to this thesis, these were not published in the literature and no details were available regarding how these enzymes catalyse their chemistry. Using sequence analysis and my crystal structure, amino acids were identified that may be involved in catalysis. These residues were mutated and their participation in catalysis was assessed by kinetic experiments. During the course of this thesis, a structural analysis and an ^{1}H NMR based kinetic analysis of *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase was published (Pelissier *et al.*, 2014). From this data, the authors proposed that the catalytic mechanism employed by *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes involve deprotonation and reprotonation of the substrate by a catalytically essential lysine residue that acts both as a Brønsted base on one stereoisomer and an acid on the other. Although this type of mechanism is common mechanistic strategy for carbohydrate epimerases (Tanner, 2002), my critical analysis of their structures suggests that this mechanism is unlikely. For this to occur, the substrate and active site would have to reorient, for which there is no evidence.

Alternatively, it is more likely that catalysis occurs *via* the formation of a Schiff base or a proton displacement mechanism mediated by the substrate. Both of which would utilise the catalytically essential lysine residue that is ideally positioned relative to the site of epimerisation on the substrate. The formation of a Schiff base was ruled out because treatment of the enzyme with sodium borohydride did not result in inactivation, which

would reduce a Schiff base. Instead, I propose that catalysis occurs *via* a proton displacement mechanism mediated by the substrate. In this mechanism, a proton is abstracted from the C2 site by the catalytically essential Lys63 residue and reprotonation of the opposite face is mediated by the substrate hydroxyl on the C5 position, which is ideally placed to do so and is activated by Glu180. This mechanism is consistent with the crystallographic data.

3.4 References

Allard, S. T., Giraud, M. F. & Naismith, J. H. (2001). Epimerases: structure, function and mechanism. *Cellular and Molecular Life Sciences, 58*, 1650-1665.

Almagro-Moreno, S. & Boyd, E. F. (2009). Sialic acid catabolism confers a competitive advantage to pathogenic *Vibrio cholerae* in the mouse intestine. *Infection and Immunity, 77*, 3807-3816.

Bohm, G., Muhr, R. & Jaenicke, R. (1992). Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Engineering, 5*, 191-195.

Boyce, J. M. (1992). Methicillin-resistant Staphylococcus aureus in hospitals and long-term care facilities: microbiology, epidemiology, and preventive measures. *Infect Control Hosp Epidemiol, 13*, 725-737.

Chambers, H. F. & Deleo, F. R. (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nature Reviews Microbiology, 7*, 629-641.

Chang, D. E., Smalley, D. J., Tucker, D. L., Leatham, M. P., Norris, W. E., Stevenson, S. J., Anderson, A. B., Grissom, J. E., Laux, D. C., Cohen, P. S. & Conway, T. (2004). Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 7427-7432.

Cornish-Bowden, A. (2004). *Fundamentals of enzyme kinetics*. London, Portland Press Ltd.

Exnowitz, F., Meyer, B. & Hackl, T. (2012). NMR for direct determination of K_m and V_{max} of enzyme reactions based on the Lambert W function-analysis of progress curves. *Biochimica et Biophysica Acta, 1824*, 443-449.

Ferrero, M. A., Martinez-Blanco, H., Lopez-Velasco, F. F., Ezquerro-Saenz, C., Navasa, N., Lozano, S. & Rodriguez-Aparicio, L. B. (2007). Purification and characterisation of GlcNAc-6-P 2-epimerase from *Escherichia coli* K92. *Acta Biochimica Polonica, 54*, 387-399.

Grundmann, H., Aires-de-Sousa, M., Boyce, J. & Tiemersma, E. (2006). Emergence and resurgence of meticillin-resistant *Staphylococcus aureus* as a public-health threat. *Lancet, 368*, 874-885.

Hinderlich, S., Stasche, R., Zeitler, R. & Reutter, W. (1997). A bifunctional enzyme catalyzes the first two steps in *N*-acetylneuraminic acid biosynthesis of rat liver -

purification and characterisation of UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase. *Journal of Biological Chemistry, 272*, 24313-24318.

Jeong, H. G., Oh, M. H., Kim, B. S., Lee, M. Y., Han, H. J. & Choi, S. H. (2009). The capability of catabolic utilisation of *N*-acetylneuraminic acid, a sialic acid, is essential for *Vibrio vulnificus* pathogenesis. *Infection and Immunity, 77*, 3209-3217.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*, 2947-2948.

Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Cambridge, The Royal Society of Chemistry.

Lawrence, M., Barbosa, J., Smith, B., Hall, N., Pilling, P., Ooi, H. & Marcuccio, S. (1997). Structure and mechanism of a sub-family of enzymes related to *N*-acetylneuraminate lyase. *Journal of Molecular Biology, 266*, 381-399.

Lebowitz, J., Lewis, M. S. & Schuck, P. (2002). Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Science, 11*, 2067-2079.

Lee, Y. C., Wu, H. M., Chang, Y. N., Wang, W. C. & Hsu, W. H. (2007). The central cavity from the (alpha/alpha)$_6$ barrel structure of *Anabaena* sp. CH1 *N*-acetyl-D-glucosamine 2-epimerase contains two key histidine residues for reversible conversion. *Journal of Molecular Biology, 367*, 895-908.

Luchansky, S. J., Yarema, K. J., Takahashi, S. & Bertozzi, C. R. (2003). GlcNAc 2-epimerase can serve a catabolic role in sialic acid metabolism. *Journal of Biological Chemistry, 278*, 8035-8042.

Maitra, P. K. (1971). Glucose and fructose metabolism in a phosphoglucoisomeraseless mutant of *Saccharomyces cerevisiae*. *Journal of Bacteriology, 107*, 759-769.

McDonald, M., Hurse, A. & Sim, K. N. (1981). Methicillin-resistant *Staphylococcus aureus* bacteraemia. *Medical Journal of Australia, 2*, 191-194.

Mirwaldt, C., Korndorfer, I. & Huber, R. (1995). The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. *Journal of Molecular Biology, 246*, 227-239.

Murkin, A. S., Chou, W. K., Wakarchuk, W. W. & Tanner, M. E. (2004). Identification and mechanism of a bacterial hydrolysing UDP-*N*-acetylglucosamine 2-epimerase. *Biochemistry, 43*, 14290-14298.

North, R. A., Kessans, S. A., Griffin, M. D., Watson, A. J., Fairbanks, A. J. & Dobson, R. C. (2014). Cloning, expression, purification, crystallisation and preliminary X-ray diffraction analysis of *N*-acetylmannosamine-6-phosphate 2-epimerase from methicillin-resistant *Staphylococcus aureus*. *Acta Crystallographica Section F, Structural Biology and Crystallisation Communications, 70*, 650-655.

Olson, M. E., King, J. M., Yahr, T. L. & Horswill, A. R. (2013). Sialic acid catabolism in *Staphylococcus aureus*. *Journal of Bacteriology, 195*, 1779-1788.

Panlilio, A. L., Culver, D. H., Gaynes, R. P., Banerjee, S., Henderson, T. S., Tolson, J. S. & Martone, W. J. (1992). Methicillin-resistant *Staphylococcus aureus* in U.S. hospitals, 1975-1991. *Infection Control and Hospital Epidemiology, 13*, 582-586.

Pelissier, M. C., Sebban-Kreuzer, C., Guerlesquin, F., Brannigan, J. A., Bourne, Y. & Vincent, F. (2014). Structural and functional characterisation of the *Clostridium perfringens N*-acetylmannosamine-6-phosphate 2-epimerase essential for the sialic acid salvage pathway. *Journal of Biological Chemistry, 289*, 35215-35224.

Petoukhov, M. V. & Svergun, D. I. (2007). Analysis of X-ray and neutron scattering from biomacromolecular solutions. *Current opinion in structural biology, 17*, 562-571.

Pezzicoli, A., Ruggiero, P., Amerighi, F., Telford, J. L. & Soriani, M. (2012). Exogenous sialic acid transport contributes to group B streptococcus infection of mucosal surfaces. *Journal of Infectious Diseases, 206*, 924-931.

Plumbridge, J. & Vimr, E. (1999). Convergent pathways for utilisation of the amino sugars *N*-acetylglucosamine, *N*-acetylmannosamine, and *N*-acetylneuraminic acid by *Escherichia coli*. *Journal of Bacteriology, 181*, 47-54.

Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics, 40*, 191-285.

Ringenberg, M. A., Steenbergen, S. M. & Vimr, E. R. (2003). The first committed step in the biosynthesis of sialic acid by *Escherichia coli* K1 does not involve a phosphorylated *N*-acetylmannosamine intermediate. *Molecular Microbiology, 50*, 961-975.

Samuel, J. & Tanner, M. E. (2002). Mechanistic aspects of enzymatic carbohydrate epimerisation. *Nature Product Reports, 19*, 261-277.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modelling. *Biophysical Journal, 78*, 1606-1619.

Sola-Carvajal, A., Sanchez-Carron, G., Garcia-Garcia, M. I., Garcia-Carmona, F. & Sanchez-Ferrer, A. (2012). Properties of BoAGE2, a second *N*-acetyl-D-glucosamine 2-epimerase from *Bacteroides ovatus* ATCC 8483. *Biochimie, 94*, 222-230.

Stasche, R., Hinderlich, S., Weise, C., Effertz, K., Lucka, L., Moormann, P. & Reutter, W. (1997). A bifunctional enzyme catalyses the first two steps in *N*-acetylneuraminic acid biosynthesis of rat liver. Molecular cloning and functional expression of UDP-*N*-acetyl-glucosamine 2-epimerase/*N*-acetylmannosamine kinase. *Journal of Biological Chemistry, 272*, 24319-24324.

Suzuki, H., Tabata, K., Morita, E., Kawasaki, M., Kato, R., Dobson, R. C., Yoshimori, T. & Wakatsuki, S. (2014). Structural basis of the autophagy-related LC3/Atg13 LIR complex: recognition and interaction mechanism. *Structure, 22*, 47-58.

Svergun, D., Barberato, C. & Koch, M. H. J. (1995). CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography, 28*, 768-773.

Svergun, D. I. (1992). Determination of the regularisation parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography, 25*, 495-503.

Tanner, M. E. (2002). Understanding nature's strategies for enzyme-catalysed racemisation and epimerisation. *Accounts of Chemical Research, 35*, 237-246.

Tanner, M. E. (2005). The enzymes of sialic acid biosynthesis. *Bioorganic Chemistry, 33*, 216-228.

van Overtveldt, S., Verhaeghe, T., Joosten, H. J., van den Bergh, T., Beerens, K. & Desmet, T. (2015). A structural classification of carbohydrate epimerases: from mechanistic insights to practical applications. *Biotechnology Advances, 33*, 1814-1828.

Vann, W. F., Tavarez, J. J., Crowley, J., Vimr, E. & Silver, R. P. (1997). Purification and characterisation of the *Escherichia coli* K1 *neu*B gene product *N*-acetylneuraminic acid synthetase. *Glycobiology, 7*, 697-701.

Vincent, F., Davies, G. J. & Brannigan, J. A. (2005). Structure and kinetics of a monomeric glucosamine 6-phosphate deaminase: missing link of the NagB superfamily? *Journal of Biological Chemistry, 280*, 19649-19655.

Walters, D. M., Stirewalt, V. L. & Melville, S. B. (1999). Cloning, sequence, and transcriptional regulation of the operon encoding a putative *N*-acetylmannosamine-6-phosphate epimerase (*nan*E) and sialic acid lyase (*nan*A) in *Clostridium perfringens*. *Journal of Bacteriology, 181*, 4526-4532.

Wierenga, R. K. (2001). The TIM-barrel fold: a versatile framework for efficient enzymes. *Federation of European Biochemical Societies Letters, 492*, 193-198.

Zhu, A., Romero, R. & Petty, H. R. (2011). An enzymatic colourimetric assay for glucose-6-phosphate. *Analytical Biochemistry, 419*, 266-270.

Chapter Four

Towards a structural and functional understanding of bacterial sialic acid transporters

The work presented in this chapter is collaborative with Dr Rosmarie Friemann's group at the University of Gothenburg, Sweden. I visited the University of Gothenburg for a total of six months, three of which were funded by a fellowship from the European Molecular Biology Organisation (EMBO). I undertook the studies presented in this chapter with guidance from this group. Work that was independently completed by members of Dr Rosmarie Friemann's group is mentioned here. Specifically, Elin Dunevall completed the cloning, overexpression trials and detergent screening for the *Staphylococcus aureus* and *Proteus mirabilis* sodium solute symporters for sialic acid transport. Together, we developed the purification protocol for the *S. aureus* transporter, which was later adapted by Elin for the *P. mirabilis* transporter. In collaboration with Professor Jeff Abramson's group at the University of California, Los Angeles, Elin successfully crystallised the *P. mirabilis* transporter. At the Univeristy of Gothenburg, Elin Dunevall and Elin Claesson optimised these crystals and X-ray diffraction data was successfully collected. Together, we attempted to soak crystals with heavy metals, but this was unsuccessful. Instead, Weixiao Yuan Wahlgren, also from the University of Gothenburg, successfully obtained the phases by selenomethionine replacement and solved the structure. This structure is presented in Section 4.2.7 of this chapter, which I analysed, wrote and produced the figures.

4.1 Introduction

4.1.1 Sialic acid transporters

Bacterial pathogens have evolved multiple mechanisms to transport sialic acid across the cytoplasmic membrane. These include an adenosine triphosphate (ATP) binding cassette transporter (ABC), a tripartite ATP independent periplasmic transporter (TRAP), a sugar proton symporter (NanT) and a sodium solute symporter (SSS). ABC transporters utilise the energy derived from ATP hydrolysis to drive the transport of a molecule across the membrane and are classed as primary transporters (Higgins, 1992). Conversely, TRAP, NanT and SSS transporters use an electrochemical gradient to facilitate transport and are classed as secondary transporters (Kelly & Thomas, 2001). This chapter is focused on the single component NanT and SSS sialic acid transporters.

4.1.2 Sialic acid is transported using specific transporters

A functional sialic acid transporter has proven to be essential for the uptake of sialic acid in a range of human bacterial pathogens, including *Escherichia coli* (Vimr & Troy, 1985), *Haemophilus influenzae* (Mulligan *et al.*, 2012), *S. aureus* (Olson *et al.*, 2013) and *Vibrio cholerae* (Severi *et al.*, 2005). This means that without a specific sialic acid transporter, these organisms cannot acquire sialic acid from their surrounding environment. Moreover, inhibition of sialic acid transporters leads to non-virulence in animal models (Fuller *et al.*, 2000; Chang *et al.*, 2004; Severi *et al.*, 2007). Thus, developing novel molecules that target these transporters is a plausible mechanism for blocking pathogen colonisation.

Whilst most bacterial pathogens possess only one type of sialic acid transporter, there are a few exceptions that are predicted to have two types from different families. For example, *Photobacterium profundum* has an SSS and a TRAP sialic acid transporter, *Streptococcus pneumoniae* has an SSS and a putative ABC sialic acid transporter and *Salmonella enterica* has an SSS and a NanT sialic acid transporter (Severi *et al.*, 2010). It is not understood why these organisms produce more than one type of transporter, but it is possible that they import sialic acid derivatives that are known in biological contexts.

4.1.3 Sugar proton symporters for sialic acid transport

The sugar proton symporter from *E. coli*, known as 'NanT', was the very first sialic acid transporter to be discovered (Vimr & Troy, 1985). NanT is a secondary transporter, meaning that it utilises an electrochemical gradient to facilitate the energetically unfavourable transport of sialic acid against its concentration gradient. NanT belongs to the major facilitator family and co-transports sialic acid and protons in the presence of a pH gradient (Martinez *et al.*, 1995; Vimr *et al.*, 2004; Severi *et al.*, 2010; Mulligan *et al.*, 2012). Sodium, lithium or potassium ions play no part in sialic acid transport and have been demonstrated to have a slight inhibitory effect (Severi *et al.*, 2010; Mulligan *et al.*, 2012). This may be a result of an increase in the ionic strength when these ions are present (Severi *et al.*, 2010). The stoichiometry of protons required for the transport of sialic acid by NanT is yet to be confirmed.

Most major facilitator superfamily transporters consist of 12 transmembrane helices, placing both the N- and the C-termini on the cytoplasmic side of the membrane (Huang *et al.*, 2003). The transmembrane helices are split into two six-helix bundles that are connected by an extended central loop within the cytoplasm (Maiden *et al.*, 1987). Based on deoxyribonucleic acid (DNA) sequence analysis, the *E. coli* NanT sialic acid transporter is predicted to consist of 14 transmembrane helices (Figure 4.1) (Martinez *et al.*, 1995), as opposed to the usual 12 transmembrane helices found in the major facilitator superfamily (Huang *et al.*, 2003). Twelve of the transmembrane helices that make up NanT contain residues that are conserved within the major facilitator superfamily (Vimr *et al.*, 2004). However, two centrally located helices at position seven and eight (shaded black in Figure 4.1) are neither observed nor conserved in other secondary transporters (Martinez *et al.*, 1995; Vimr *et al.*, 2004). The amino acids comprising the seventh transmembrane helix are predicted to form an amphipathic helix, which may be important for the specificity of this protein towards sialic acid (Vimr *et al.*, 2004).

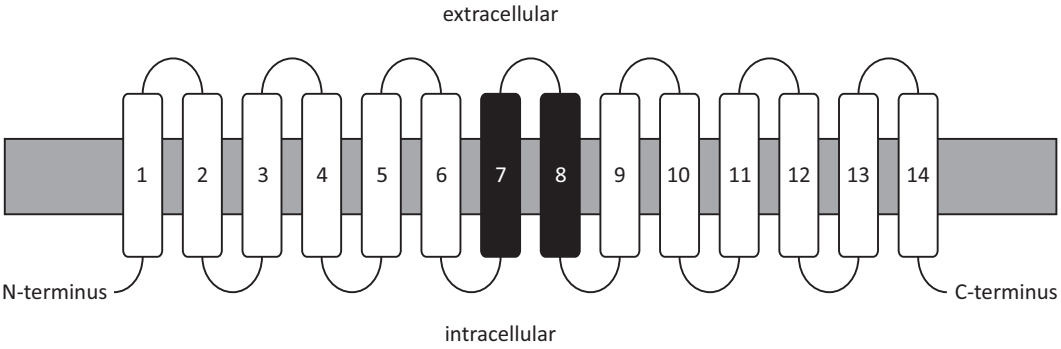


Figure 4.1. A topological model of *E. coli* NanT. The white rectangles indicate the 12 transmembrane helices that are common within major facilitator superfamily members. The central black rectangles indicate those that are unique to *E. coli* NanT [adapted from Martinez *et al.* (1995)].

Unfortunately, our understanding of sugar proton symporters for sialic acid transport at a molecular level is limited to the *E. coli* NanT (Martinez *et al.*, 1995). However, a number of crystal structures have been solved for other members of the major facilitator superfamily, the first of which includes the *E. coli* lactose permease sugar proton symporter (Abramson *et al.*, 2003) and the glycerol-3-phosphate inorganic phosphate antiporter (Huang *et al.*, 2003). The mechanism by which these transporters are proposed to move molecules across a membrane is described as the 'rocker switch' mechanism (Abramson *et al.*, 2003; Huang *et al.*, 2003). In this mechanism, the transporter alternates between an inward and outward facing conformation, where the substrate binding site is only accessible on one side of the membrane at any given time (Huang *et al.*, 2003).

4.1.4 Sodium solute symporters for sialic acid transport

A more recently discovered type of sialic acid transporter, belonging to the SSS family of secondary transporters, was found in *S. enterica* (Severi *et al.*, 2010). Orthologues of this transporter are encoded in the genome of a range of other clinically important human bacterial pathogens, including *S. aureus*, *S. pneumoniae* and *Clostridium perfringens*. Importantly, the genes encoding the SSS sialic acid transporters are widespread among both gram positive and gram negative species of bacteria, whereas other types of sialic acid transporters do not appear to be as widespread (Severi *et al.*, 2010). In contrast to the NanT sialic acid transporters that utilise a proton gradient for symport, SSS transporters depend on a sodium gradient for the movement of sialic acid across the membrane (Wright *et al.*,

2004; Severi *et al.*, 2010). It has long been established that cells maintain a low intracellular concentration of sodium ions by actively pumping them out of the cell. In combination with the negative potential of the membrane, this provides the driving force for the transport of solutes into the cell (Schultz & Curran, 1970; Faham *et al.*, 2008).

Currently, the only representative of the SSS family for which the crystal structure has been solved for is the sodium galactose transporter (SGLT) from *Vibrio parahaemolyticus* (Faham *et al.*, 2008). The structure contains 14 transmembrane helices with an extracellular facing N- and C-termini. The transmembrane helices are arranged in an inward facing conformation made up of two inverted repeats, which each contain five transmembrane helices (Figure 4.2). Despite no sequence similarity, these 10 helices display a highly similar topology to the core structure of a sodium leucine symporter (LeuT) from *Aquifex aeolicus*, belonging to the neurotransmitter sodium symporter family (Yamashita *et al.*, 2005). More recently, the crystal structure of a sodium hydantoin symporter (Mhp1) from *Microbacterium liquefaciens* (Weyand *et al.*, 2008) and a sodium glycine betaine symporter (BetP) from *Corynebacterium glutamicum* (Ressl *et al.*, 2009) have been shown to possess this same core architecture. These transporters belong to the nucleobase cation symport 1 family and the betaine/choline/carnitine transporter family, respectively. These transporters also do not share any sequence homology with one another or with the SSS family (Weyand *et al.*, 2008; Abramson & Wright, 2009). Thus, this core domain of inverted topology represents a structural hallmark for many sodium driven transporters that is not based upon sequence (Yamashita *et al.*, 2005; Weyand *et al.*, 2008; Ressl *et al.*, 2009). The additional non-conserved helices at the N- and C-termini of each respective structure may be important for function (Faham *et al.*, 2008).

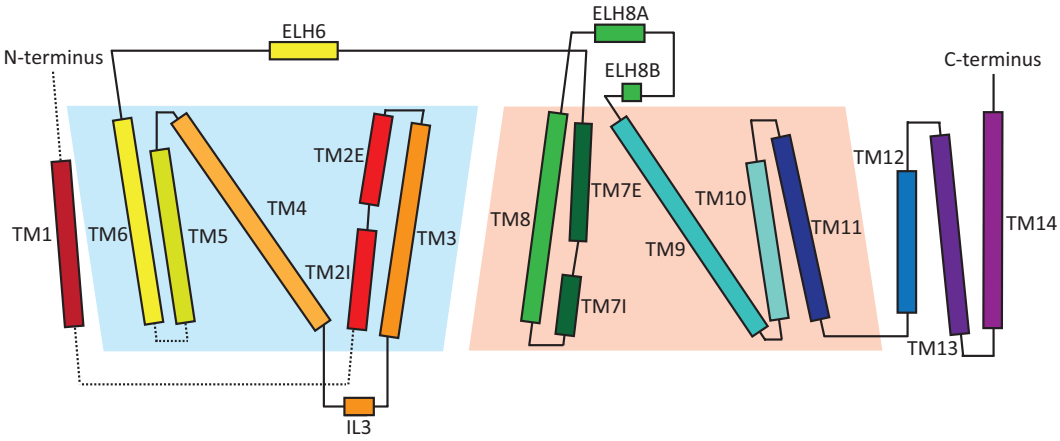


Figure 4.2. The topology derived from the structure of the sodium and galactose symporter from *V. parahaemolyticus*. The blue and red trapeziums represent the two inverted repeats made up of transmembrane helices 2 to 6 (blue) and 7 to 11 (red) [adapted from Faham *et al.* (2008)].

The stoichiometry of sodium ions required for transport by SSS transporters varies amongst family members. For some, two sodium ions are required for transport of their respective solute (Eskandari *et al.*, 1997; Mackenzie *et al.*, 1998), whereas others require only one sodium ion (Turk *et al.*, 2000). Based on a comparison with the *A. aeolicus* LeuT structure (Yamashita *et al.*, 2005), sequence analysis and mutational analysis, a putative sodium binding site was identified in the crystal structure of the *V. parahaemolyticus* SGLT. But unlike *A. aeolicus* LeuT, which possesses two sodium binding sites, a second site could not be detected in this symporter (Faham *et al.*, 2008).

4.1.5 A proposed mechanism of sodium and solute symport

Similar to the mechanism described for the major facilitator superfamily transporters, sodium symporters also share an alternating access mechanism, where the transporter switches between an inward and an outward facing conformation for transport (Faham *et al.*, 2008). Based on the observation that a common core structure is shared between *V. parahaemolyticus* SGLT, *A. aeolicus* LeuT, *M. liquefaciens* Mhp1 and *C. glutamicum* BetP and that the crystal structures reside in distinct conformations, a model for the mechanism of sodium and solute symport has been proposed (Figure 4.3) (Abramson & Wright, 2009).

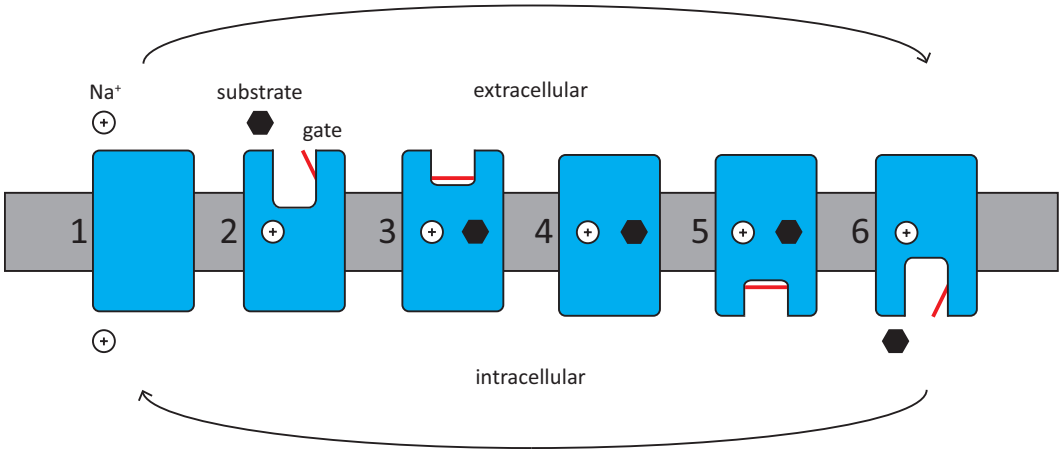


Figure 4.3. The proposed mechanism of sodium and solute symport. In this model the starting conformation has no ligands bound (position 1). Following the binding of a sodium ion/ions from the extracellular side, an external gate (red) is opened so that the structure is in an outward facing and open conformation, with a pathway that leads to the substrate binding site (position 2). Upon substrate binding, the external gate closes causing the structure to be in an outward facing, but occluded conformation (position 3). The corresponding conversion from the outward to the inward facing occluded state occurs through an intermediate state, where the substrate is completely sequestered in the core of the transporter. The external pathway closes and there is no way that the substrate can exit (position 4). Once the inward facing conformation is formed, the internal pathway begins to open (position 5). Opening of the internal gate permits the sodium ion/ions and substrate to be released into the intracellular environment (position 6). The cycle is completed and the transporter returns to the starting conformation (position 1). It is important to note that the dwell time of ligands bound at any conformation is not known [adapted from Abramson and Wright (2009)].

The structure of *M. liquefaciens* Mhp1 has a sodium ion bound and is in an outward facing and open conformation (Weyand *et al.*, 2008) (position 2, Figure 4.3). This provides evidence for a pathway that leads to the substrate binding site. The sodium and substrate bound structure of *A. aeolicus* LeuT is in an outward facing and occluded conformation, which suggests that upon substrate binding an external gate composed of hydrophobic residues is closed, which in turn occludes the substrate from the extracellular environment (Yamashita *et al.*, 2005) (position 3, Figure 4.3). Yet, a pathway above the closed gate at the extracellular milieu remains intact. Next, it is proposed that the transporter will transition from an outward to an inward facing conformation through an intermediate state. This transition was discovered in the structure of *C. glutamicum* BetP (Ressl *et al.*, 2009) (position 4, Figure 4.3). Finally, evidence for an inward facing and occluded conformation (position 5, Figure 4.3) is provided by the *V. parahaemolyticus* SGLT structure (Faham *et*

al., 2008). Here, the substrate is bound in the centre of the core and occluded from the outside solutions by hydrophobic residues that form an intracellular gate. To date, there are no structures of an inward facing and open conformation devoid of substrate (position 6, Figure 4.3), so the nature of this transition cannot yet be inferred (Abramson & Wright, 2009; Mulligan *et al.*, 2011).

4.1.6 Overview of this chapter

To gain an understanding of how sialic acid is transported across the membrane of bacterial pathogens, this chapter explores three sialic acid transporter proteins, belonging to two distinct gene families. Data are presented for the NanT sialic acid transporter from *Yersinia pestis* and the SSS sialic acid transporters from *S. aureus* and *P. mirabilis*. Importantly, sialic acid transporters have proven to be essential for the survival of many harmful bacterial pathogens (Vimr & Troy, 1985; Chang *et al.*, 2004; Severi *et al.*, 2005; Severi *et al.*, 2007; Mulligan *et al.*, 2012; Olson *et al.*, 2013). However, there are no structural data to explain how sialic acid transporters mediate transport across the membrane or how specificity is achieved. Thus, the overarching goal of this chapter is to gain structural data for these sialic acid transporters, which will underpin the design of future antibiotic drugs.

This chapter presents the overexpression, solubilisation, purification and possible oligomeric states of these transporters; a significant step towards a structural and functional understanding of these difficult to work with proteins. In addition, a potential protein-protein interaction between the *S. aureus* SSS sialic acid transporter and *N*-acetylneuraminate lyase is investigated. Excitingly, the *P. mirabilis* SSS sialic acid transporter was successfully crystallised and the structure was solved, which has provided insight into how this transporter mediates the movement of sialic acid across the membrane. Because sialic acid transporters are antibiotic drug targets, elucidating the structural basis for symport will greatly facilitate our understanding of the diseases caused by harmful bacterial pathogens and aid in the development of novel antibiotic drugs.

4.2 Results and discussion

4.2.1 Cloning and overexpression of the *Y. pestis* NanT, *S. aureus* SSS and *P. mirabilis* SSS sialic acid transporters

It is well recognised in the literature that the overexpression of membrane proteins is notoriously difficult (Gordon *et al.*, 2008). When overexpressed in large quantities, the membrane protein will often become toxic to the overexpression host. This toxicity is most likely a result of saturating the membrane protein biogenesis/protein secretion machinery (Wagner *et al.*, 2008; Schlegel *et al.*, 2012). Tracking the production of folded and functional membrane protein throughout overexpression is difficult. Thus, the cloning and overexpression systems employed for such proteins need to be carefully thought out. A detailed account of the cloning and overexpression strategy used for the *Y. pestis* NanT, *P. mirabilis* SSS and *S. aureus* SSS sialic acid transporters is presented.

The *E. coli* BL21(DE3) strain in combination with T7-ribonucleic acid (RNA) polymerase promoter (T7*lac*) based plasmids is the most commonly used platform to drive the overexpression of proteins (Schlegel *et al.*, 2012). This is a bacteriophage T7-based pET/T7-RNA polymerase overexpression system, in which overexpression of the target protein is governed by T7-RNA polymerase. The isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible *lac*UV5 promoter governs overexpression of this T7-RNA polymerase (Wanner *et al.*, 1977). The rationale behind this system is that the more mRNA that is produced, the more protein that can be overexpressed (Schlegel *et al.*, 2012).

Recently, it has been shown that mutations in the *lac*UV5 promoter improve the overexpression of membrane proteins (Wagner *et al.*, 2008). Because this promoter governs the overexpression of T7-RNA polymerase, which in turn governs the overexpression of a target protein, mutations in the *lac*UV5 promoter result in a simultaneous decrease in the production of T7-RNA polymerase and target protein. As a consequence, the risk of saturating the membrane protein biogenesis/protein secretion capacity upon overexpression is diminished (Schlegel *et al.*, 2012). Thus, a modified strain of BL21(DE3), known as Lemo21(DE3), has been developed in which production of T7-

RNA polymerase can be controlled by co-expression of its inhibitor, T7 lysozyme (Figure 4.4). The gene encoding T7 lysozyme is located on a separate plasmid known as pLemo. The rhamnose (*rha*BAD) promoter on this pLemo plasmid governs overexpression of T7 lysozyme. This promoter is titratable, meaning that the amount of rhamnose that is added correlates with the amount of T7 lysozyme that is overexpressed (Giacalone *et al.*, 2006). For instance, the more rhamnose that is added, the more T7 lysozyme that is overexpressed and the more T7-RNA polymerase that is inhibited, meaning that there is less T7-RNA polymerase available to induce the T7*lac* promoter for overexpression of the target protein and vice versa. As a consequence, Lemo21(DE3) is tunable and allows for conditions to be optimised for stable membrane protein overexpression (Schlegel *et al.*, 2012).

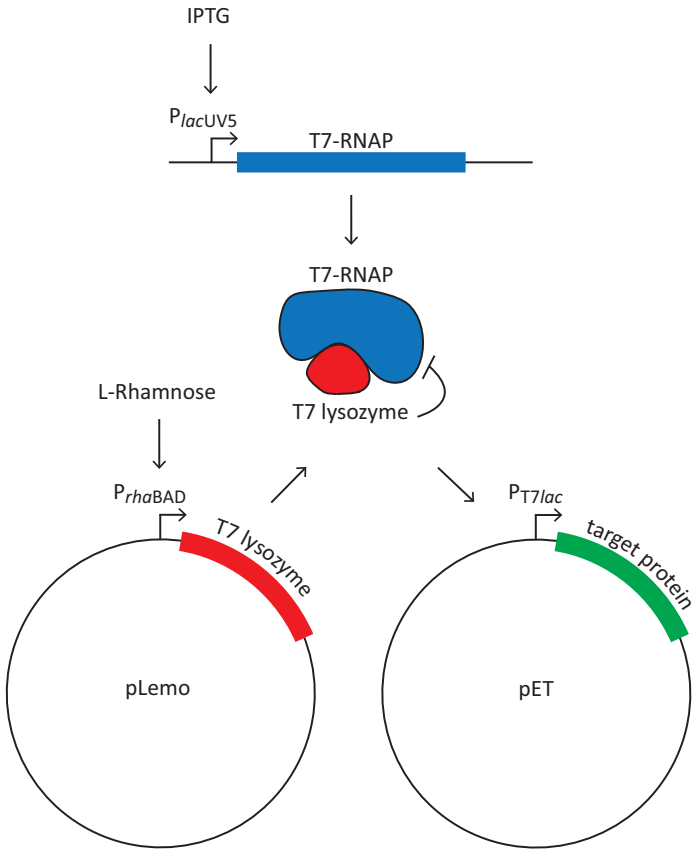


Figure 4.4. A schematic representation of Lemo21(DE3). The IPTG-inducible *lac*UV5 promoter governs overexpression of the chromosomally located gene that encodes T7-RNA polymerase. The gene encoding T7 lysozyme, an inhibitor of T7-RNA polymerase, is located on the pLemo plasmid. The titratable rhamnose promoter (*rha*BAD) governs its overexpression. The gene encoding the target protein is located on a pET vector and the T7*lac* promoter governs its overexpression [adapted from Schlegel *et al*. (2012)].

Previously, it was shown that C-terminal green fluorescent protein (GFP) fusion tags could be used to monitor the overexpression of bacterial membrane proteins by measuring whole cell fluorescence (Drew *et al.*, 2001). Membrane protein overexpression levels in the cytoplasmic membrane of *E. coli* can be estimated with a detection limit as low as 10 μg of protein per litre of bacterial culture (Drew *et al.*, 2006; Hjelm *et al.*, 2013). In addition, GFP fluorescence can also be detected by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), with a detection limit of less than 5 ng of GFP (Drew *et al.*, 2006). Importantly, the GFP is only fluorescent if the overexpressed protein is in a soluble form and correctly folded (Hjelm *et al.*, 2013). Thus, this extremely stable GFP moiety can be easily visualised, allowing for both the amount and integrity of the target membrane protein to be monitored at any stage during the overexpression procedure.

The gene encoding the *Y. pestis* NanT sialic acid transporter (Appendix) was commercially synthesised with *Xho*I and *Eco*RI restriction sites in the pMA-RQ cloning plasmid. The gene was excised from pMA-RQ and subcloned into the pWarf(−) expression plasmid to produce pWarf(−)NanT (Chapter Six, Section 6.4.6-6.4.8). Dr Rosmarie Friemann's group supplied the genes encoding the *S. aureus* and *P. mirabilis* SSS sialic acid transporters (Appendix), also in the pWarf(−) expression plasmid. The pWarf(−) expression plasmid contains a C-terminal GFP fusion tag, allowing for the overexpression, detergent solubilisation and purification to be monitored through fluorescence (Drew *et al.*, 2006; Hsieh *et al.*, 2010). These constructs were transformed into *E. coli* Lemo21(DE3) bacterial cells and overexpression trials were performed in terrific broth (TB) medium (Chapter Six, Section 6.6.1). TB is an enriched medium that has been shown to achieve higher overexpression levels for membrane proteins (Hsieh *et al.*, 2010).

For overexpression of the *Y. pestis* NanT sialic acid transporter, a range of L-rhamnose concentrations (0-2000 μM) was tested, at various induction times and at either 25 or 30 °C. In all conditions, the concentration of IPTG was kept constant, as this is not an important variable for overexpression in this system (Drew *et al.*, 2006; Hsieh *et al.*, 2010). Both whole cell fluorescence (Chapter Six, Section 6.6.4) and in-gel fluorescence by SDS-PAGE (Chapter Six, Section 6.5.2) were utilised to analyse which overexpression condition is the most appropriate and to assess the integrity and quantification of overexpression. As shown in Figure 4.5 A, an overnight induction with 500 μM of L-rhamnose had the highest whole cell fluorescence at both 25 and 30 °C. In addition, an

overnight induction with 250 µM L-rhamnose at 25 °C and 1000 µM L-rhamnose at 30 °C displayed the next highest whole cell fluorescence. All four of these overexpression conditions exhibited a relative fluorescent unit (RFU) above 15, which is the minimum overexpression criterion necessary for the isolation of a target protein in sufficient quantity for structural and functional studies (Hsieh *et al.*, 2010). Although the NanT-GFP fusion fusion protein has a predicted molecular weight of 84.7 kDa, it appears at approximately 50 kDa when subjected to in-gel fluorescence (Figure 4.5 B). This is often the case for membrane proteins when subjected to SDS-PAGE, but the reasons behind this phenomenon are not conclusively explained. It has been suggested that the binding of detergent molecules to the protein of interest may explain anomalous SDS-PAGE behaviour (Rath *et al.*, 2009). The most intense fluorescent bands in Figure 4.5 B correspond to 500 µM of L-rhamnose at both 25 and 30 °C (lane 2 and lane 3, respectively), supporting the whole cell fluorescence data. Although a small amount of free GFP can be detected in the gel, the NanT-GFP fusion protein is present in adequate quantities for structural and functional studies.

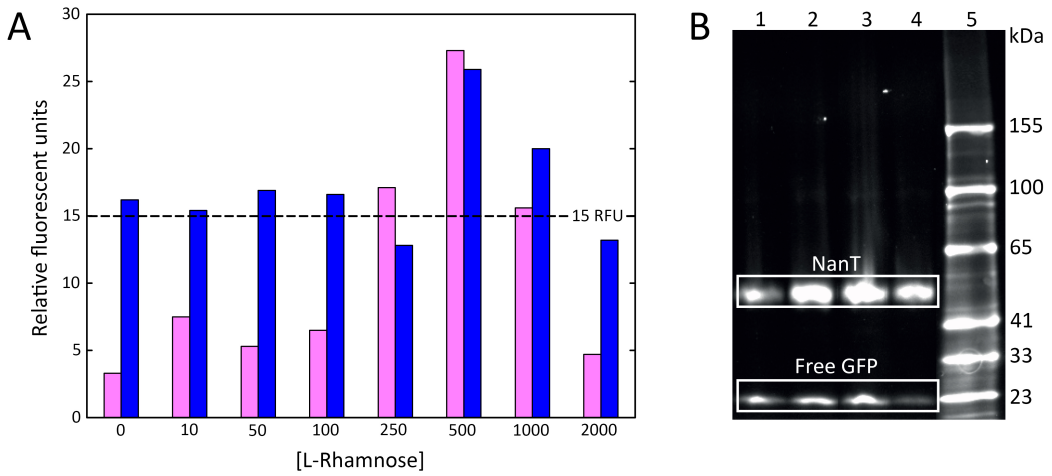


Figure 4.5. Whole cell fluorescence and in-gel fluorescence of the *Y. pestis* NanT sialic acid transporter in different overexpression conditions. (A) Whole cell fluorescence reported as RFU. The *Y. pestis* NanT sialic acid transporter was overexpressed at either 25 (magenta) or 30 °C (blue) overnight, across a range of L-rhamnose concentrations. The horizontal dashed line indicates 15 RFU, which is the minimum required for overexpression of a suitable amount of protein. **(B)** In-gel fluorescence of the *Y. pestis* NanT sialic acid transporter from the four best overexpression conditions determined by whole cell fluorescence. Lane 1, 250 µM L-rhamnose at 25 °C; lane 2, 500 µM L-rhamnose at 25 °C; lane 3, 500 µM L-rhamnose at 30 °C; lane 4, 1000 µM L-rhamnose at 30 °C; lane 5, protein ladder (kDa).

From these results combined, the overexpression conditions selected for the *Y. pestis* NanT sialic acid transporter in Lemo21(DE3) bacterial cells consisted of an overnight induction with 500 μM of L-rhamnose, at a temperature of 25 °C. The conditions required for overexpression of the *S. aureus* SSS and *P. mirabilis* SSS sialic acid transporters were previously established as 100 μM and 250 μM of L-rhamnose, respectively. Culturing was also performed at a temperature of 25 °C and an overnight induction time. Following determination of the appropriate overexpression conditions, cells were cultured and protein was overexpressed at a larger scale. Cells were harvested and lysed (Chapter Six, Section 6.6.2-6.6.3) and then the membranes were collected (Chapter Six, Section 6.6.5).

4.2.2 Detergent screening of the *Y. pestis* NanT sialic acid transporter

The development of membrane protein-GFP fusion methodology to monitor overexpression has been expanded to identify detergents that are suitable for both purification and crystallisation studies. Often, membrane protein crystallisation is only achieved in the presence of a detergent that maintains the protein in a monodisperse and stable state (Ostermeier & Michel, 1997), where the best detergent is usually one that mimics the proteins' native membrane environment (Hsieh *et al.*, 2010). Thus, successful crystallisation of a membrane protein is highly dependent upon the detergent that has been selected. Traditionally, detergent screening is a major obstacle for membrane protein crystallisation, as the methods require milligram quantities of already purified protein and are both time and resource intensive (Kawate & Gouaux, 2006; Hsieh *et al.*, 2010).

Fluorescence detection size exclusion chromatography (FSEC) has emerged as an efficient method for identifying an ideal detergent prior to extensive purification trials (Kawate & Gouaux, 2006). Because FSEC can exploit the unique fluorescence signal of GFP, crude samples can be subjected to this technique. This strategy involves solubilising the crude membrane collection with a variety of detergents and directly applying them to a size exclusion column equilibrated in a buffer containing the respective detergent (Chapter Six, Section 6.6.7). Here, the elution profile allows the monodispersity and stability of the membrane protein-GFP fusion to be monitored by fluorescence spectroscopy (Kawate & Gouaux, 2006; Hsieh *et al.*, 2010). A monodisperse and stable protein will typically yield a

single and symmetrical peak, while a polydisperse and unstable protein will reveal multiple asymmetric peaks (Ricker & Sandoval, 1996; Kawate & Gouaux, 2006).

Using FSEC, the *Y. pestis* NanT sialic acid transporter was subjected to a panel of six detergents, including n-dodecyl-N,N-dimethylamine-N-oxide (LDAO), n-octyl-β-D-glucopyranoside (NG), 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), n-decyl-β-D-maltoside (DM), n-dodecyl-β-D-maltoside (DDM) and 2,2-didecylpropane-1,3-bis-β-D-maltopyranoside (MNG), all of which are suitable for membrane protein structural and functional studies (Wiener, 2004; Prive, 2007; Chae *et al.*, 2010). The *Y. pestis* NanT sialic acid transporter was unstable and formed aggregates in LDAO and NG complexes, identified by a single peak in the void volume of the FSEC profile. In CHAPS, the *Y. pestis* NanT sialic acid transporter did not form significant aggregates, but the elution profile was polydisperse (Figure 4.6 A). In DDM and MNG, the *Y. pestis* NanT sialic acid transporter displayed asymmetric and monodisperse peaks in the FSEC profile (Figure 4.6 B and C, respectively). However, there were no significant aggregates identified in the void volume of the column with this detergent.

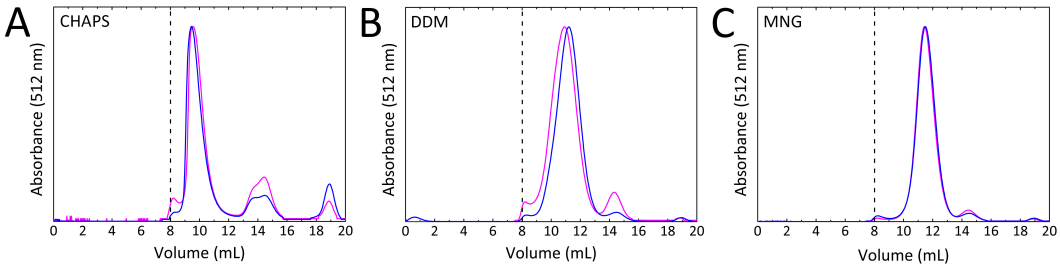


Figure 4.6. Detergent screening of the *Y. pestis* NanT sialic acid transporter using FSEC. NanT was solubilised in CHAPS **(A)**, DDM **(B)** and MNG **(C)** and then subjected to analysis by FSEC. The traces in blue were obtained immediately after solubilisation, while the traces in magenta were obtained 48 h post solubilisation. The fluorescence traces were normalised between 0 and 1 for each detergent and time point. The void volume of the column is indicated with vertical dashes.

To assess whether the *Y. pestis* NanT sialic acid transporter and detergent complexes with DDM, MNG and CHAPS were stable over time, the complexes were again evaluated by FSEC after 48 hours (Hsieh *et al.*, 2010). After 48 hours, the FSEC profile of each of the *Y. pestis* NanT sialic acid transporter and detergent complexes remained similar. With CHAPS and DDM, the *Y. pestis* NanT sialic acid transporter showed a very small increase

in the amount of aggregate after 48 hours, depicted by a slight increase in the peak at the void volume of the column. In addition, a small amount of protein degradation was observed for the *Y. pestis* NanT sialic acid transporter with CHAPS, as indicated by an increase in the size of the second polydisperse peak seen in the chromatogram. The peak for the *Y. pestis* NanT sialic acid transporter with DDM shifted slightly after 48 hours, but overall, the protein and detergent complex remained stable and monodisperse over time. The *Y. pestis* NanT sialic acid transporter with MNG gave a similar elution profile after 48 hours, meaning that this complex remained stable and monodisperse. In conclusion, both DDM and MNG are suitable detergents for the *Y. pestis* NanT sialic acid transporter, resulting in stable and monodisperse protein over time. DDM was selected for purification and additional experiments conducted with the *Y. pestis* NanT sialic acid transporter in this chapter. However, MNG should also be assessed for the purification and crystallisation of this protein. For the *S. aureus* and *P. mirabilis* SSS sialic acid transporters, DM and DDM were respectively selected as the best detergents.

4.2.3 Purification of the *Y. pestis* NanT, *S. aureus* SSS and *P. mirabilis* SSS sialic acid transporters

4.2.3.1 Immobilised metal affinity chromatography and protease cleavage

A 6 × Histidine tag (His-tag) follows the C-terminal GFP fusion protein overexpressed with each transporter, meaning that each of these proteins could be subjected to immobilised metal affinity chromatography (IMAC) for purification (Chapter Six, Section 6.6.8.3). Importantly, the fluorescence associated with the GFP fusion protein could be followed throughout each purification step. This allowed for the integrity and stability of the membrane protein to be monitored throughout purification trials. The purity and stability of each transporter following purification was assessed by SDS-PAGE and in-gel fluorescence. The purity of the *Y. pestis* NanT, *S. aureus* SSS and *P. mirabilis* SSS sialic acid transporters following purification by IMAC is depicted in Figure 4.7 A, B and C, respectively. For the *S. aureus* and *P. mirabilis* SSS sialic acid transporters, a band that is approximately double the size of the monomeric band is visualised by in-gel fluorescence; this may in fact be a higher oligomeric species.

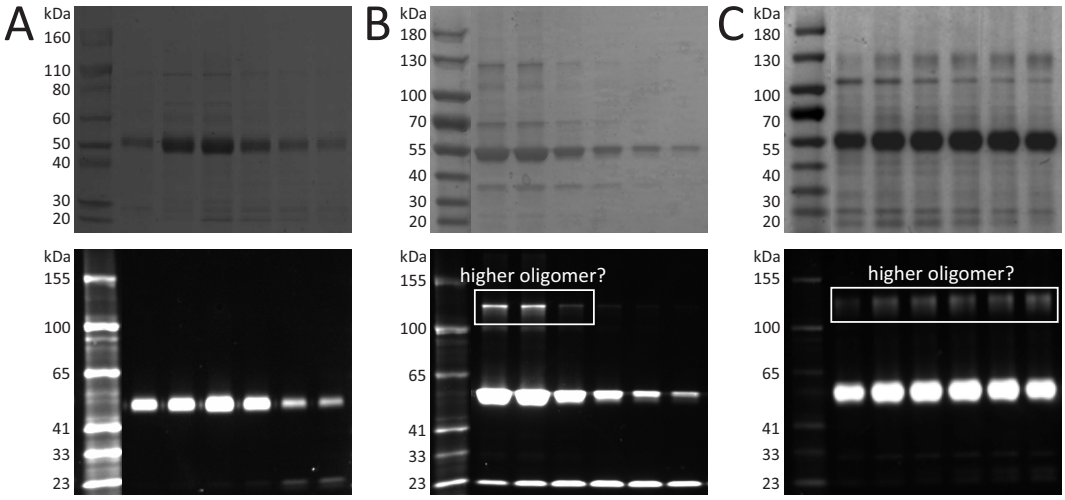


Figure 4.7. SDS-PAGE and in-gel fluorescence analysis following IMAC. Fractions corresponding to elution of the *Y. pestis* NanT **(A)**, *S. aureus* SSS **(B)** and *P. mirabilis* SSS **(C)** sialic acid transporters using IMAC. SDS-PAGE analysis is shown at the top, while the corresponding in-gel fluorescence analysis is shown at the bottom. The protein ladder (kDa) is located on the left of each SDS-PAGE gel.

Because artificial fusion tags can sometimes have unfavourable effects in both structural and functional studies (Hsieh *et al.*, 2010), the human rhinovirus 3C (HRV3C) protease cleavage site was engineered between the respective membrane protein and the C-terminal GFP fusion moiety with a His-tag. This allows for the GFP fusion moiety and His-tag to be removed. Following cleavage (Chapter Six, Section 6.6.8.5), the transporters were recovered by gravity flow chromatography (Chapter Six, Section 6.6.8.6), before being subjected to size exclusion chromatography.

4.2.3.2 Size exclusion chromatography

Size exclusion chromatography (Chapter Six, Section 6.6.8.7) was employed as the final purification step with overall yields presented in Table 4.1. Size exclusion chromatography of the *Y. pestis* NanT sialic acid transporter gave a 280 nm absorbance peak in the void volume of the column, indicating that considerable aggregation had occurred (Figure 4.8 A). However, the peak corresponding to the elution of non-aggregated NanT is sharp and symmetrical in shape, suggesting that the non-aggregated sample is homogenous. Interestingly, the *S. aureus* SSS sialic acid transporter elutes across two distinct peaks, indicative that the transporter may exist in more than one oligomeric state (Figure 4.8 B).

Importantly, pooling the second peak corresponding to a smaller species (fractions at 11.5-12.5 mL) and subjecting it to another round of size exclusion chromatography generates a similar profile. Consistent with IMAC, SDS-PAGE analysis displays a band that is approximately double the size of the monomeric band, providing further evidence for a higher oligomeric state. The *P. mirabilis* SSS sialic acid transporter elutes as a single peak, with a small shoulder at the beginning of elution making it slightly asymmetric (Figure 4.8 C). Although not as distinctive as the double peak observed for the *S. aureus* SSS sialic acid transporter, SDS-PAGE analysis suggests that a higher oligomeric state might also be present for this transporter. The *S. aureus* SSS and *P. mirabilis* SSS sialic acid transporters do not display any significant aggregation.

Table 4.1. Purification of each sialic acid transporter. The culture size, membrane yield and protein yield is indicated as well as the yield of protein from each litre of culture.

Construct	Culture (L)	Membrane (g)	Protein (mg)	Yield (mg/L of culture)
Y. pestis NanT	4	7.3	0.8	0.2
S. aureus SSS	3	5.1	1.9	0.6
P. mirabilis SSS	12	20.7	10.6	0.9

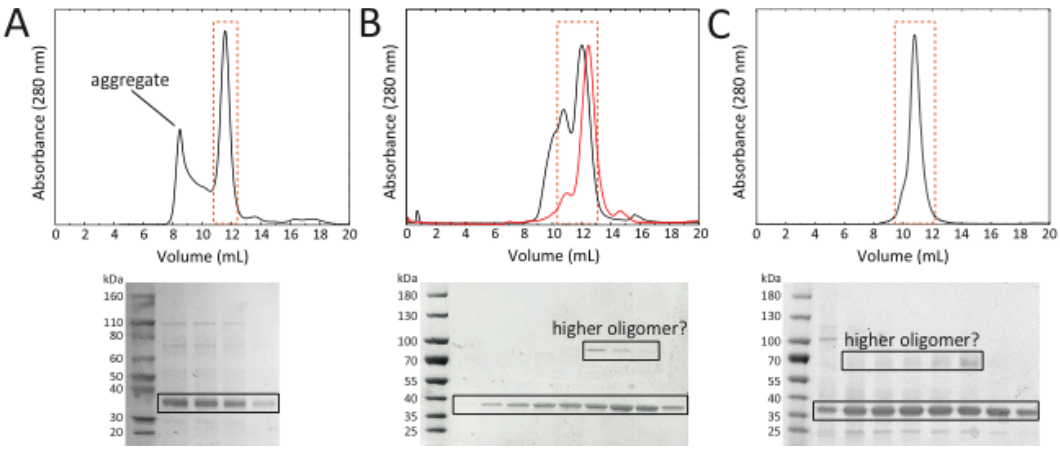


Figure 4.8. The purity of each sialic acid transporter determined by size exclusion chromatography and SDS-PAGE analysis. Chromatograms for size exclusion chromatography (top) and SDS-PAGE analysis (bottom) for the *Y. pestis* NanT **(A)**, *S. aureus* SSS **(B)** and the *P. mirabilis* SSS **(C)** sialic acid transporters. For the *S. aureus* SSS sialic acid transporter, the red trace corresponds to re-injection of the second peak in black (fractions at 11.5-12.5 mL), which has been normalised to the original data. The areas of the 280 nm absorbance peaks that were analysed by SDS-PAGE are highlighted with an orange box. The protein ladder is located on the far left of each SDS-PAGE gel (kDa).

4.2.4 The oligomeric structure of the *Y. pestis* NanT, *S. aureus* SSS and *P. mirabilis* SSS sialic acid transporters

Sedimentation velocity experiments by analytical ultracentrifugation were used to assess the oligomeric structure of the *Y. pestis* NanT, *S. aureus* SSS and the *P. mirabilis* SSS sialic acid transporters solvated in a detergent solution. Although challenging, this technique has been successful for membrane transporters in the past (Ebel, 2011). For the *S. aureus* SSS and *P. mirabilis* SSS sialic acid transporters, evidence for a higher oligomeric state was seen during purification; analytical ultracentrifugation can be used to further confirm this notion. A limitation of this work is that the detergent used for each protein is not accounted for in the density and viscosity calculations for the respective buffers. However, by fitting the data to a sedimentation coefficient distribution [c*(s)*] model, it is possible to qualitatively determine if the protein is a single species, or whether it self-associates in the detergent micelles.

4.2.4.1 The solution structure of the Y. pestis *NanT sialic acid transporter*

The c*(s)* distribution of the *Y. pestis* NanT sialic acid transporter shows a dominant single and symmetrical peak at a standardised sedimentation coefficient relative to water at 20 °C ($s_{20,w}$) of 6.6, suggesting that the transporter is largely a single species in solution (Figure 4.9 A). Importantly, most detergents do not absorb at 280 nm (Slotboom *et al.*, 2008; Le Roy *et al.*, 2015) and the peak seen corresponds to the protein. The distribution also shows evidence of contaminants from the purification, or aggregated protein, but this only represents 9% of the signal. SDS-PAGE analysis following size exclusion chromatography revealed that the sample is largely pure (Figure 4.8 A), but minor bands were still visible that could account for the presence of contaminants in the sample. However, considerable aggregation was also seen during size exclusion chromatography and it is more likely that the protein again began to aggregate during data collection in the analytical ultracentrifuge. The raw data and the fit generated by the c*(s)* model is plotted and the residuals are nearly randomly distributed around zero, indicating a good fit to the data (Figure 4.9 B) (Lebowitz *et al.*, 2002).

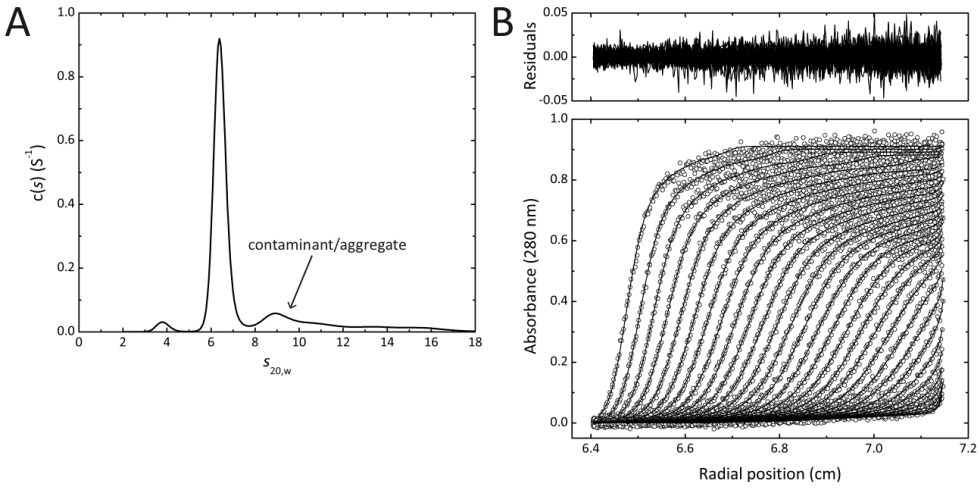


Figure 4.9. Sedimentation velocity analysis of the *Y. pestis* NanT sialic acid transporter. The experiment was conducted at 0.5 mg/mL (8.9 μM) of freshly purified protein in 50 mM 2-amino-2-hydroxymethyl-propane-1,3-diol (Tris), pH 8.0, 150 mM sodium chloride, 6% glycerol, 0.0174% DDM. Data were collected at 280 nm, 50 000 rpm and 4 °C. SEDNTERP was used to calculate the partial specific volume of the *Y. pestis* NanT sialic acid transporter **(**0.7505 g/ml), solvent density (1.0248 g/ml) and viscosity (0.0193 poise) (Laue *et al.*, 1992). In these calculations, the detergent (DDM) was ignored. Data were fitted to a c(*s*) model at a resolution of 300 and a confidence level of 0.95 using SEDFIT (Schuck, 2000). **(A)** The c(*s*) model plotted as a function of the $s_{20,w}$. Data were fitted with an *s* value ranging between 1 and 10 S. **(B)** Absorbance at 280 nm plotted as a function of radial position (cm). The raw data are represented as open symbols (o) and overlaid with the non-linear least squares best fit. The residuals for this fit are shown above.

4.2.4.2 The solution structure of the S. aureus *SSS sialic acid transporter*

The c(*s*) distribution for *S. aureus* SSS sialic acid transporter resulted in two peaks with $s_{20,w}$ values of 6.4 and 8.8 S at 0.125 mg/mL (2.2 μM) (Figure 4.10 A). At 0.25 mg/mL (4.4 μM) these peaks were also observed as well as an additional peak at 11.4 S, suggesting that this transporter is forming a self-association in solution (Ebel, 2011; Le Roy *et al.*, 2015). This data is consistent with the SDS-PAGE and in-gel fluorescence analysis following purification (Figures 4.7 and 4.8 B), where a species approximately twice the mass of the monomer is observed. Moreover, it is also consistent with the size exclusion chromatography data, which demonstrated that pooling and re-injecting the initial larger peak also gave rise to two peaks (Figure 4.8 B). In addition, because the size exclusion chromatography data resolved into these distinct peaks, the self-association may

have relatively slow rate constants for association (k_{on}) and dissociation (k_{off}). Although this analysis is only qualitative and the exact oligomeric structure cannot be determined, I propose that depending on the concentration, the *S. aureus* SSS sialic acid transporter is self-associating to form a monomer, dimer, trimer and so on in solution. The raw data and the fit generated by the c*(s)* model is plotted for the experiment conducted at 2.2 µM (Figure 4.10 B). The residuals are nearly randomly distributed around zero at both concentrations, indicating a good fit to the data (Figure 4.10 C) (Lebowitz *et al.*, 2002).

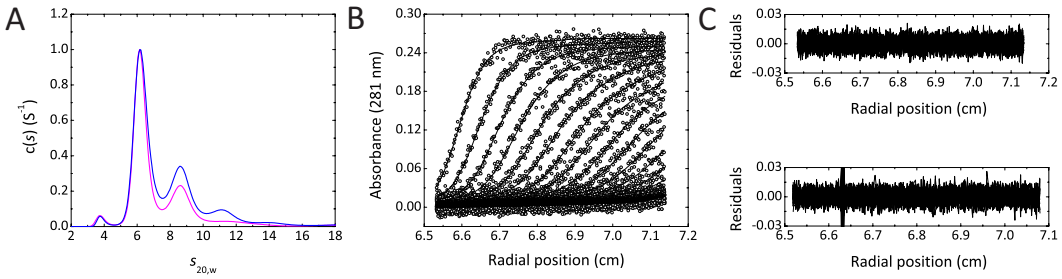
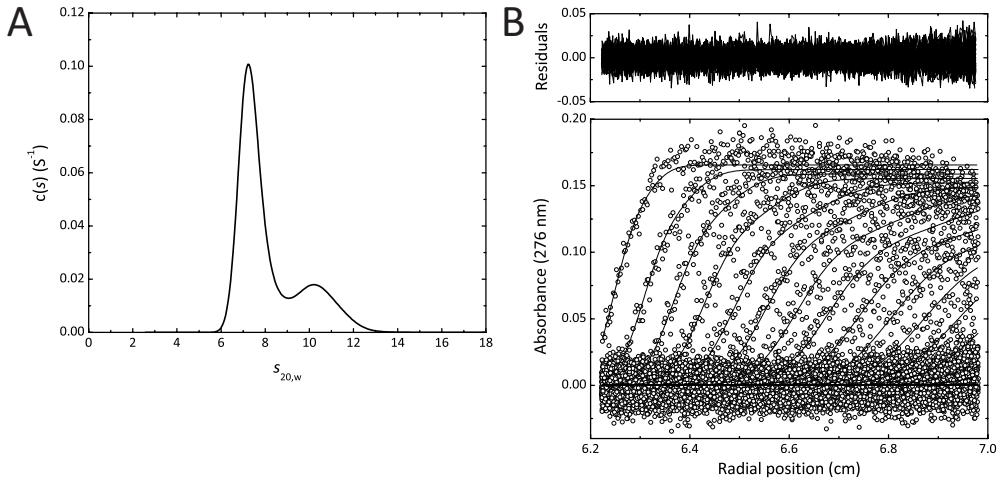


Figure 4.10. Sedimentation velocity analysis of the *S. aureus* SSS sialic acid transporter at two concentrations. The experiments were conducted at 2.2 µM (magenta) and 4.4 µM (blue) of freshly purified protein in 50 mM Tris, pH 8.0, 150 mM sodium chloride, 0.174% DM. Data were collected at 281 nm, 50 000 rpm and 12 °C. SEDNTERP was used to calculate the partial specific volume of the *S. aureus* SSS sialic acid transporter (0.7568 g/ml), solvent density (1.0071 g/ml) and viscosity (0.0152 poise) (Laue *et al.*, 1992). In these calculations, the detergent (DM) was ignored. Data were fitted to a c(*s*) model at a resolution of 300 and a confidence level of 0.68 using SEDFIT (Schuck, 2000). **(A)** The c(*s*) model plotted as a function of the $s_{20,w}$. Data were fitted with an *s* value ranging between 1 and 15 S. Data were normalised between 0 and 1. **(B)** Absorbance at 281 nm plotted as a function of radial position (cm) for the experiment conducted at 2.2 µM. The raw data are represented as open symbols (o) and overlaid with the non-linear least squares best fit. **(C)** The residuals for the data collected at 2.2 µM are shown at the top and the residuals for the data collected at 4.4 µM fit are shown at the bottom.

4.2.4.3 The solution structure of the P. mirabilis *SSS sialic acid transporter*

Similar to the *S. aureus* SSS sialic acid transporter, the c(*s*) distribution for the *P. mirabilis* SSS sialic acid transporter resulted in two peaks with $s_{20,w}$ values of 7.5 and 10.6 (Figure 4.11 A), suggesting that more than one oligomeric state is present in solution. I propose that the *P. mirabilis* SSS sialic acid transporter is self-associating to form a monomer and a dimer in solution. For this protein, multiple concentrations are yet to be tested to further confirm a self-association. Although SDS-PAGE and in-gel fluorescence analysis also



suggest that a higher oligomeric state is present (Figures 4.7 and 4.8 C), the size exclusion chromatography data (Figure 4.8 C) does not show two distinct peaks like the *S. aureus* SSS sialic acid transporter. This may suggest that if in a self-association, the k_{on} and k_{off} rates are fast for the *P. mirabilis* SSS sialic acid transporter. The raw data and the fit generated by the c*(s)* model is plotted and the residuals are nearly randomly distributed around zero, indicating a good fit to the data (Figure 4.11 B) (Lebowitz *et al.*, 2002).

Figure 4.11. Sedimentation velocity analysis of the *P. mirabilis* SSS sialic acid transporter. The experiment was conducted at 0.125 mg/mL (2.3 µM) of freshly purified protein in 50 mM Tris, pH 8.0, 150 mM sodium chloride, 0.0174% DDM. Data were collected at 276 nm, 50 000 rpm and 12 °C. SEDNTERP was used to calculate the partial specific volume of the *S. aureus* SSS sialic acid transporter (0.7551 g/ml), solvent density (1.0071 g/ml) and viscosity (0.0127 poise) (Laue *et al.*, 1992). In these calculations, the detergent (DDM) was ignored. Data were fitted to c(*s*) model at a resolution of 300 and a confidence level of 0.68 using SEDFIT (Schuck, 2000). **(A)** The c(*s*) model plotted as a function of the $s_{20,w}$. Data were fitted with an s value ranging between 1 and 15 S. **(B)** Absorbance at 276 nm plotted as a function of radial position (cm). The raw data are represented as open symbols (o) and overlaid with the non-linear least squares best fit. The residuals for this fit are also shown above.

4.2.5 A potential protein-protein interaction

It has long been established that proteins involved in the same metabolic pathway and more importantly the successive enzymes of a pathway, may be involved in protein-protein interactions with one another (Durek & Walther, 2008). Interestingly, the genes encoding the *S. aureus* SSS sialic acid transporter and *N*-acetylneuraminate lyase, the first and

committed enzyme involved in the degradation of sialic acid, are co-transcribed. Moreover, as described in Chapter Two, *N*-acetylneuraminate lyase is tetrameric in structure and the four monomers associate by leaving a large pore in the central cavity of the tetramer, with the active sites oriented on the inward face of the pore. Thus, I hypothesised that a protein-protein interaction may occur between the SSS sialic acid transporter and *N*-acetylneuraminate lyase enzyme in *S. aureus*. To test this hypothesis, analytical ultracentrifugation was employed, which is a powerful tool for characterisation of protein-protein interactions (Rivas *et al.*, 1999). Sedimentation velocity analysis was performed with the *S. aureus* SSS sialic acid transporter and methicillin-resistant *S. aureus* (MRSA) *N*-acetylneuraminate lyase (Chapter Two) separately and combined. The data were fitted to a c(*s*) distribution (Figure 4.12).

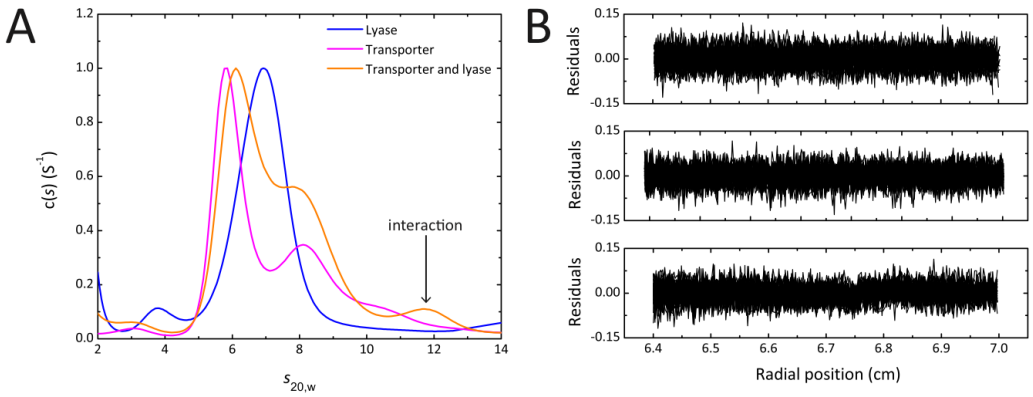


Figure 4.12. Sedimentation velocity analysis of the *P. mirabilis* SSS sialic acid transporter and MRSA *N*-acetylneuraminate lyase, to assay a possible protein-protein interaction. The experiment was conducted at 0.5 mg/mL (8.8 µM) of the *S. aureus* SSS sialic acid transporter in one cell, 0.5 mg/mL (15.1 µM) of MRSA *N*-acetylneuraminate lyase in a second cell and a combined final concentration of 0.5 mg/mL (11.9 µM) of both proteins in a third cell. Proteins were freshly purified in 50 mM Tris, pH 8.0, 150 mM sodium chloride, 0.0174% DM. Data were collected at 289 nm, 50 000 rpm and 12 °C. SEDNTERP was used to calculate the partial specific volume of the *S. aureus* SSS sialic acid transporter (0.7568 g/ml), MRSA *N*-acetylneuraminate lyase (0.7384 g/ml) and combined (0.7492 g/ml). SEDNTERP was also used to calculate the solvent density (1.0243 g/ml) and viscosity (0.0152 poise) (Laue *et al.*, 1992). In these calculations, the detergent (DM) was ignored. **(A)** The c(*s*) model plotted as a function of the $s_{20,w}$. Data were fitted with an s value ranging between 1 and 12 S. Data were normalised between 0 and 1. **(B)** The residuals for the MRSA *N*-acetylneuraminate lyase data (top), the *S. aureus* SSS sialic acid transporter data (middle) and the interaction data (bottom) are shown.

The c(*s*) distribution for the *S. aureus* SSS sialic acid transporter reveals two distinct peaks with $s_{20,w}$ values of 6.1 and 8.3. MRSA *N*-acetylneuraminate lyase reveals a dominant species with a $s_{20,w}$ of 7.5. In addition, a smaller peak with a $s_{20,w}$ value of 4.2 is apparent; this smaller peak is not seen for MRSA *N*-acetylneuraminate lyase in Chapter Two, Section 2.2.2.2, and could be a result of detergent destabilising the tetrameric structure. Excitingly, the c(*s*) distribution for both the *S. aureus* SSS sialic acid transporter and MRSA *N*-acetylneuraminate lyase combined reveals three peaks. Two with $s_{20,w}$ values of 6.4 and 8.5, respectively, which are consistent with merged c(*s*) distributions that were observed for the proteins alone. A third peak is observed with a $s_{20,w}$ value of 11.9, which is not accounted for in the *c*(*s*) distributions of these proteins alone. This provides evidence for a protein-protein interaction between one of the *S. aureus* SSS sialic acid transporter species and MRSA *N*-acetylneuraminate lyase. I propose that this interaction would occur at the intracellular face of the membrane and allow for the immediate degradation of sialic acid upon entering the cell (Figure 4.13).

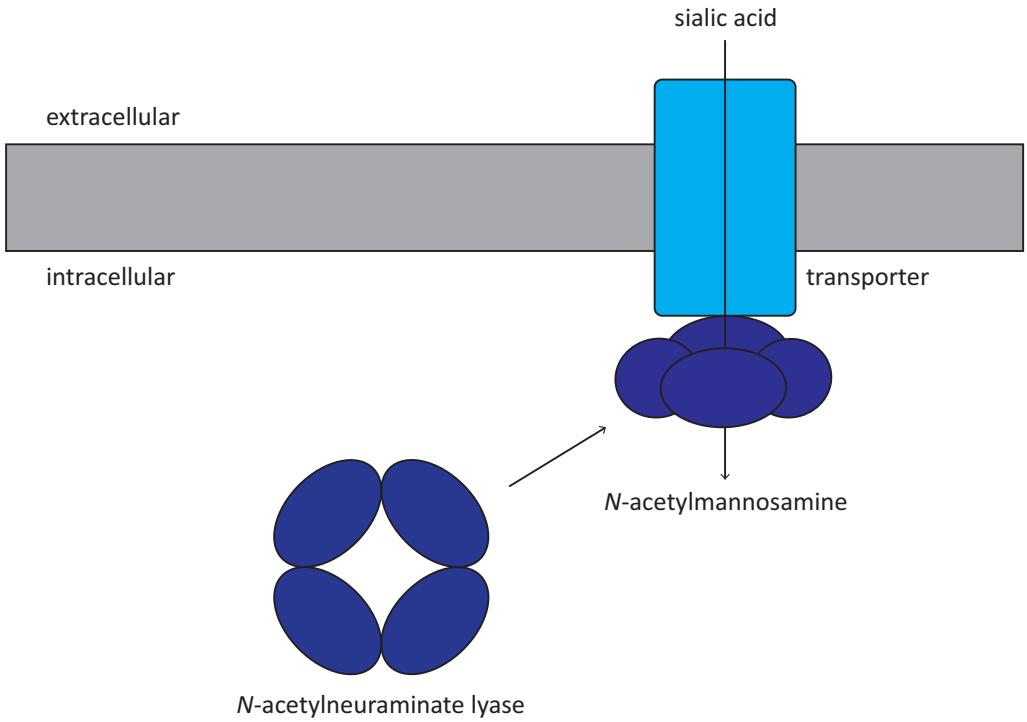


Figure 4.13. The proposed protein-protein interaction between the *S. aureus* SSS sialic acid transporter and the *N*-acetylneuraminate lyase enzyme. Based on the observation that the genes encoding these two proteins are co-transcribed in *S. aureus*, the shape of *N*-acetylneuraminate lyase and the sedimentation velocity data, an interaction between these proteins may occur on the intracellular side.

4.2.6 Crystallisation studies

Crystallisation studies were performed using both vapour diffusion and lipidic cubic phase methods. For the *Y. pestis* NanT sialic acid transporter, no crystals were obtained in lipidic cubic phase, but unusual crystals were obtained in the MemGold™ vapour diffusion screen. These crystals took at least two weeks to form at 20 °C. Interestingly, they appeared spherical in shape and were birefringent; a result of fine needles growing in all directions from an initial nucleation event. Crystals from conditions D7 (100 mM potassium chloride, 100 mM Tris, pH 8.5, 39% v/v polyethylene glycol 400) and E4 (20 mM Tris, pH 7.5; 22% v/v polyethylene glycol 500 monomethyl ether) diffracted to approximately 18 Å (Figure 4.14 A and B, respectively). These crystallisation conditions should be optimised in the future. In addition, because the *Y. pestis* NanT sialic acid transporter is also stable in MNG, purification and crystallisation with this detergent should also be tested. Unfortunately, no crystals were obtained for the *S. aureus* SSS sialic acid transporter in both vapour diffusion and lipidic cubic phase methods. Because crystallisation of a membrane protein is highly dependent upon the detergent that is being used (Hsieh *et al.*, 2010), other detergents should be trialled.

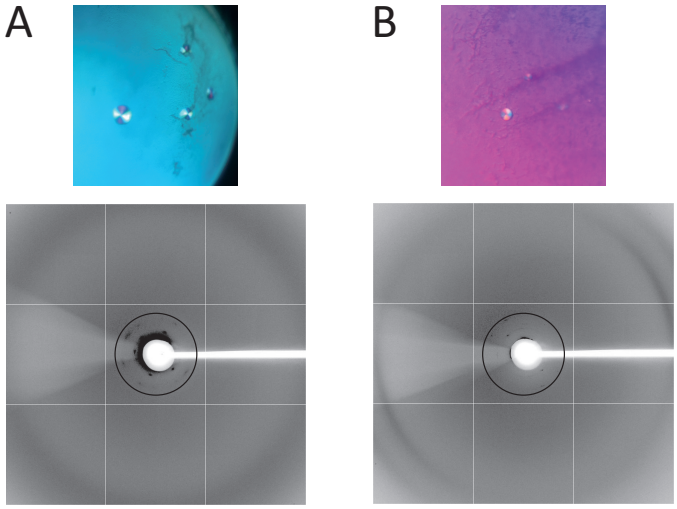


Figure 4.14. Crystals and X-ray diffraction of the *Y. pestis* NanT sialic acid transporter. (A) Crystals and diffraction from condition D7 of the MemGold™ screen. **(B)** Crystals and diffraction from condition E4 of the MemGold™ screen. The black ring indicates diffraction to approximately 18 Å in both A and B.

The *P. mirabilis* SSS sialic acid transporter was successfully crystallised using vapour diffusion methods. These crystals were optimised and an X-ray diffraction data set was collected to 2.1 Å. Molecular replacement was attempted with the only other structural representative of the SSS family, the *V. parahaemolyticus* sodium and galactose symporter (PDB entry 3dh4, 26% identity) (Faham *et al.*, 2008), but the phases could not be obtained in this way. Instead, the phases were successfully obtained by selenomethionine replacement. An X-ray diffraction dataset was collected to 1.95 Å and the structure could then be solved.

4.2.7 The structure of the *P. mirabilis* SSS sialic acid transporter

The structure of the *P. mirabilis* SSS sialic acid transporter was solved in complex with sialic acid and sodium ions bound to the protein, which was added to the crystallisation conditions to stabilise the protein and to provide molecular information on ligand binding. The structure was solved in the space group *C*2 and the unit cell parameters are a =130.6, b = 98.1, c = 54.8, α = 90.0, β = 92.2, γ = 90.0. The structure and refinement statistics are shown in Table 4.2.

Table 4.2. Structure and refinement statistics for the *P. mirabilis* SSS sialic acid transporter. The structure was solved in the space group *C*2 and there is one monomer in the asymmetric unit.

Statistics	*P. mirabilis* SSS sialic acid transporter
residual factor (R_{factor}) (%)	20.1
free R_{factor} (R_{free}) (%)	24.5
number of atoms	
protein	3702
water	242
ligands	60
root-mean-square deviation (rmsd)	
bonds (Å)	0.017
angles (°)	1.082
average B factors (Å^2)	
protein	29.7
water	44.0
ligands	40.1
Ramachandran plot, residues (%)	
favoured region	97.25
allowed region	2.33
disallowed region	0.42

4.2.7.1 The crystal lattice demonstrates how oligomerisation may occur

The asymmetric unit contains one monomer, which makes contacts with other monomers at three distinct interfaces in the crystal lattice. The monomers either side of the interfaces *A* and *B* are arranged in the opposite orientation to one another, whereas the monomers either side of interface *C* are arranged in the same orientation (Figure 4.15). As evidenced during purification and analytical ultracentrifugation, the *P. mirabilis* SSS sialic acid transporter may form a self-associated higher oligomer. Oligomerisation is likely to occur between monomers at interface *C*, which buries approximately 500 Å (approximately 3%) of the surface accessible area. Although the exact oligomeric structure of this transporter cannot be confirmed, this interface could enable dimerisation between monomers, or even higher order oligomeric structures. Moreover, whether such oligomers form in the cell membrane is unknown and it is possible that the oligomerisation seen in the solution based experiments are an artefact of the experimental conditions.

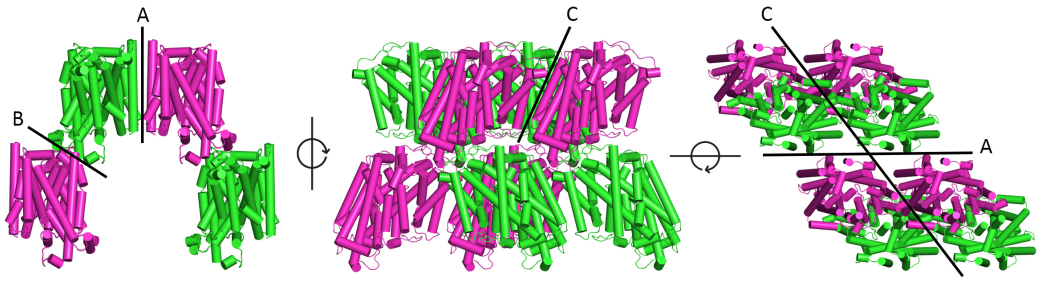


Figure 4.15. Crystal packing of the *P. mirabilis* SSS sialic acid transporter. The monomer of the *P. mirabilis* SSS sialic acid transporter interacts with other monomers across three distinct interfaces. These interfaces are labelled *A*, *B* and *C*, respectively. A self-association may be enabled across interface *C*.

If oligomerisation does occur, it is unclear what role it plays in the function of the *P. mirabilis* SSS sialic acid transporter. The interface is not extensive and it is largely made up of hydrophobic interactions, including contacts with a single DDM detergent molecule (Figure 4.16). Although this interface does not play a direct role in sialic acid or sodium ion interactions (Section 4.2.7.3), dimerisation may influence the positioning of the helices involved in these interactions and their ligand binding sites.

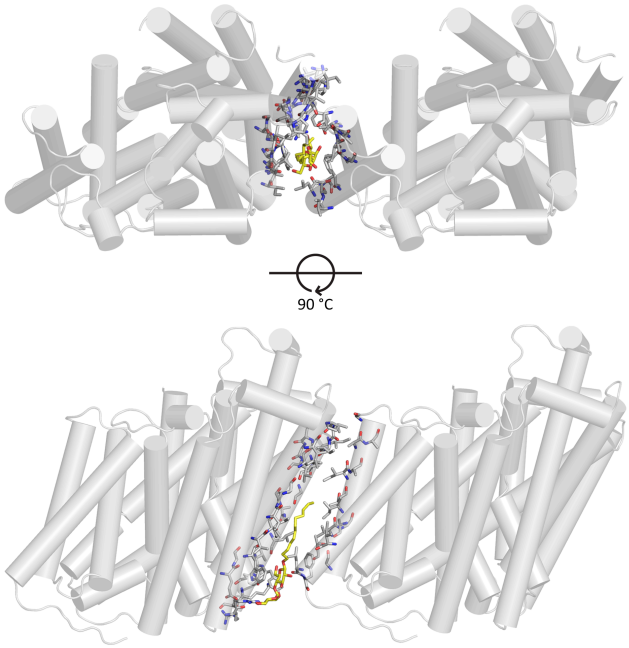


Figure 4.16. Oligomerisation of the *P. mirabilis* SSS sialic acid transporter. Dimerisation, or even higher order oligomers, may occur across interface C, which is made up of predominantly hydrophobic residues (grey sticks) and contacts with a DDM detergent molecule (yellow).

4.2.7.2 The overall structural topology

The structure contains 13 transmembrane helices (Tm1 to Tm13) with an N-terminus located in the periplasm and a C-terminus located in the cytoplasm. As depicted in Figure 4.17 A, the structural core is formed from two inverted repeats of five transmembrane helices each (Tm2 to Tm6 and Tm7 to Tm11). This topological arrangement is similar to that of the *V. parahaemolyticus* SGLT (Faham *et al.*, 2008), *A. aeolicus* LeuT (Yamashita *et al.*, 2005), *M. liquefaciens* Mhp1 (Weyand *et al.*, 2008) and *C. glutamicum* BetP (Ressl *et al.*, 2009) sodium symporters, which do not share sequence similarities. The overall structure of the *P. mirabilis* SSS sialic acid transporter is composed of five transmembrane helices that contribute interactions for ligand selectivity (Tm2, Tm3, Tm4, Tm6 and Tm9) (Figure 4.17 B), while the remaining helices are probably important for stabilisation of the structure. Numerous transmembrane helices (Tm2, Tm3, Tm5, Tm7 and Tm11) display a break in the helical structure at positions that are within the membrane bilayer. It has previously been noted that discontinuous helices are often important for function and have mechanistic implications for a range of co-transporters (Faham *et al.*, 2008).

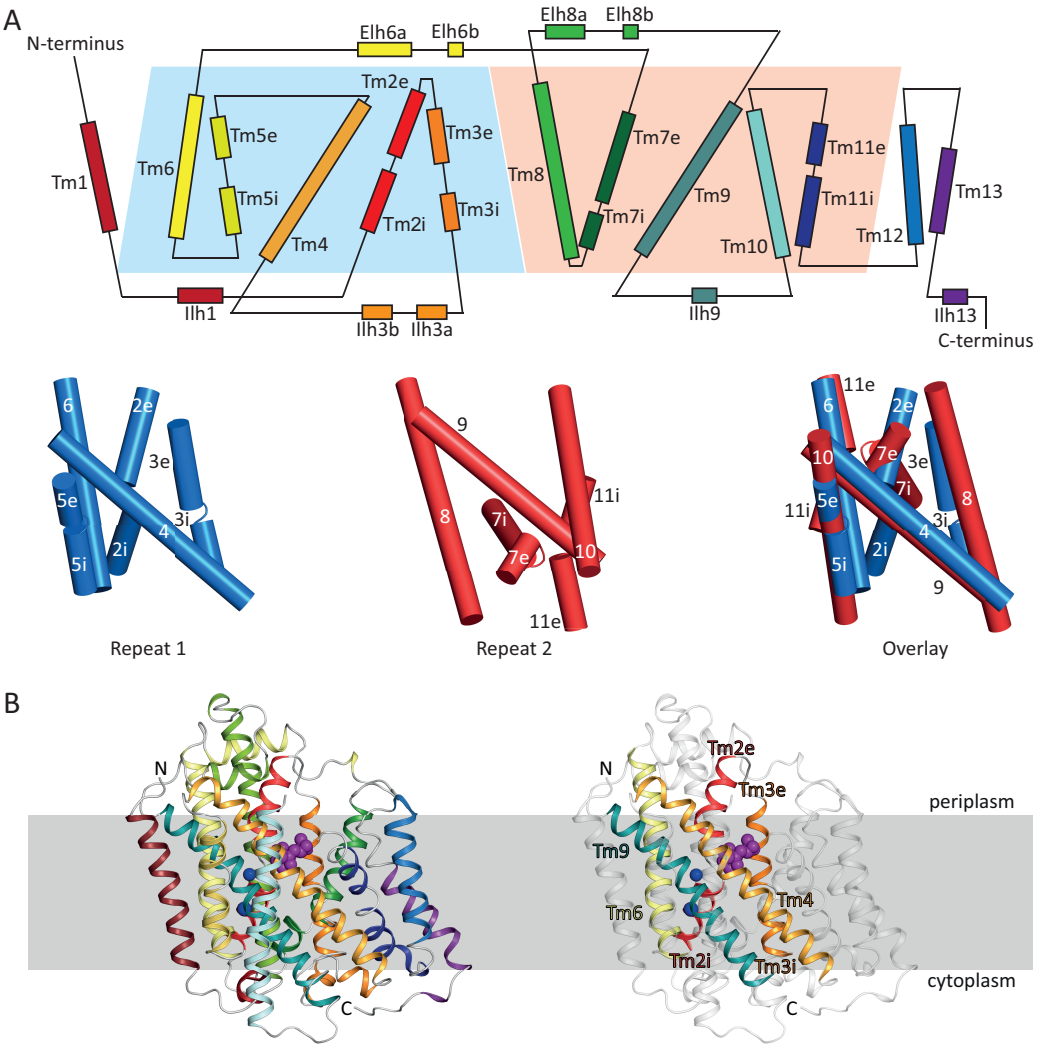


Figure 4.17. The structure of the *P. mirabilis SSS* sialic acid transporter. (A) The topology of the transporter represents an inverted repeat made up of five transmembrane helices in each. The blue and red trapeziums represent the inverted topology of Tm2 to Tm6 and Tm7 to Tm11, respectively. The inverted repeats are related by an apparent two-fold symmetry around an axis through the centre of the membrane plane. A superposition of Tm2 to Tm6 (blue) with Tm7 to Tm11 (red) yields an r.m.s.d. of 6.6 Å for 75 Cα atoms. **(B)** A side-view of the transporter in the membrane plane is shown on the left. While the transmembrane helices involved in ligand binding are depicted in colour on the right. Sialic acid is shown as purple spheres and sodium ions are shown as blue spheres.

4.2.7.3 The sialic acid and sodium binding sites

The sialic acid binding site in the *P. mirabilis* SSS sialic acid transporter forms side chain interactions with residues located on Tm2i, Tm2e, Tm3e and Tm4. Specifically, sialic acid forms hydrogen bonds with Thr58, Ser60, Thr63, Gln82 and Arg135 located on these central helices (Figure 4.18 A). Interestingly, the stoichiometry of sodium ions required for transport by SSS transporters varies amongst family members. For example, the structures of *A. aeolicus* LeuT (Yamashita *et al.*, 2005) and *C. glutamicum* BetP (Ressl *et al.*, 2009) elucidated two sodium binding sites for these transporters, whereas the structures of *V. parahaemolyticus* SGLT (Faham *et al.*, 2008) and *M. liquefaciens* Mhp1 (Weyand *et al.*, 2008) possess only one sodium binding site. The *P. mirabilis* SSS sialic acid transporter structure comprises two sodium binding sites (Figure 4.18 B), one of which forms hydrogen bonds with Ala56, Leu59, Ala339, Ser342 and Ser343 and the other forms hydrogen bonds with Thr57, Asp182, Ser342, Ser345 and Ser346. Importantly, the residues involved in binding sialic acid and sodium ions are highly conserved with SSS sialic acid transporters from a range of other organisms (Figure 4.19). Thus, it is likely that SSS sialic acid transporters utilise two sodium ions for the transport of sialic acid across the membrane.

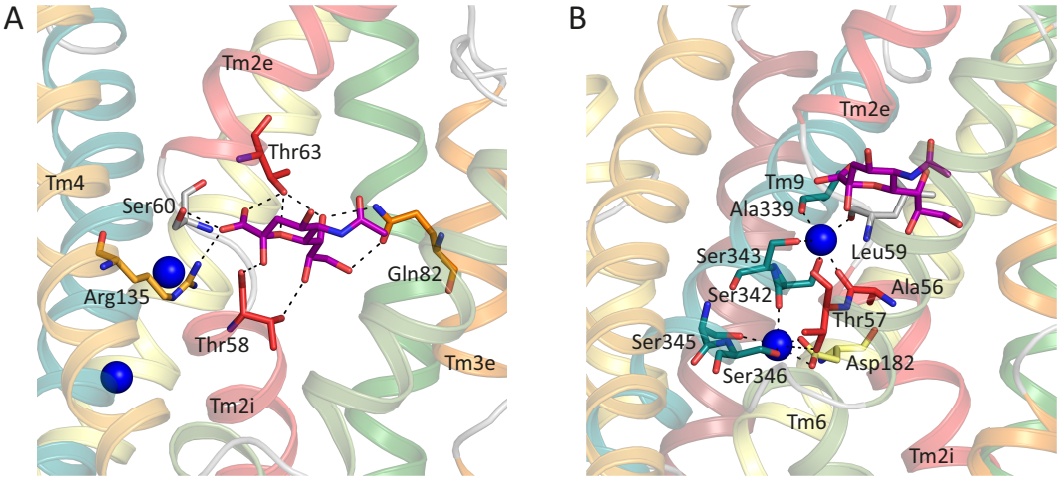


Figure 4.18. The ligand binding sites of the *P. mirabilis* SSS sialic acid transporter. (A) Sialic acid (purple sticks) forms hydrogen bonds with Thr58, Thr63, Ser60, Gln82 and Arg135. **(B)** One sodium ion (blue spheres) forms hydrogen bonds with Ala56, Leu59, Ala339, Ser342 and Ser343 while the other forms hydrogen bonds with Thr57, Asp182, Ser342, Ser345 and Ser346.

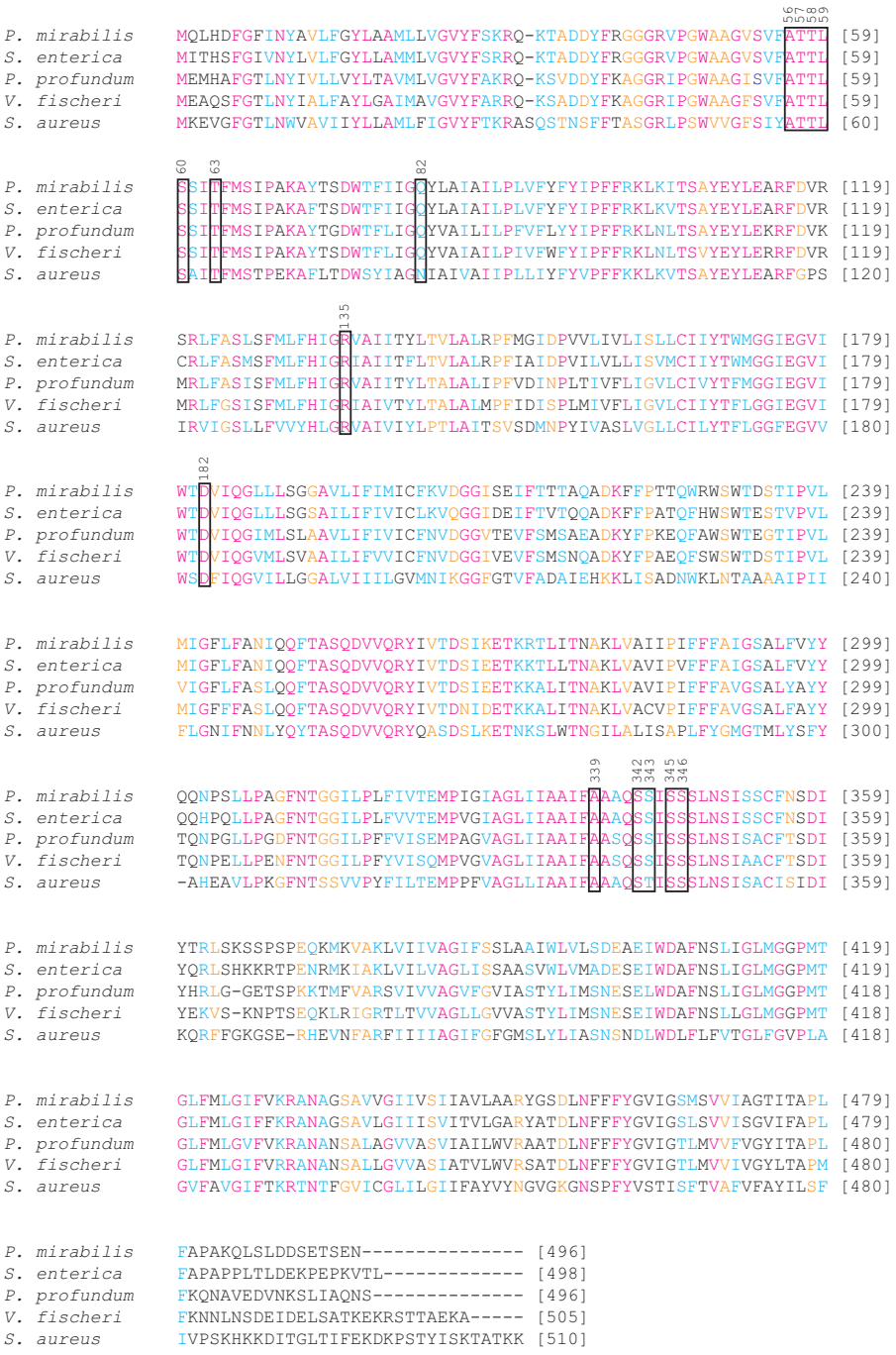


Figure 4.19. Sequence alignment of SSS sialic acid transporters from five species of bacteria. Species include *P. mirabilis*, *S. enterica*, *P. profundum*, *Vibrio fischeri* and *S. aureus*. Strictly conserved residues are highlighted in magenta, conservation between residues with strongly similar properties are shown in blue and conservation between residues with weakly similar properties are shown in orange. Residues implicated in sialic acid and sodium binding in the *P. mirabilis* SSS sialic acid transporter are highlighted with a black box and numbered according to this transporter.

4.2.7.4 A novel outward facing and open conformation

Importantly, our structure of the *P. mirabilis* SSS sialic acid transporter is in an outward facing and open conformation, with both sodium and sialic acid bound. The sialic acid sits within a cavity that is open on the periplasmic side, which can be seen at the top of the structure (Figure 4.20). The *A. aeolicus* LeuT sodium and leucine symporter is also in an outward facing conformation with both sodium and substrate bound (Yamashita *et al.*, 2005). However, this structure is in an occluded conformation and the substrate is sandwiched between residues that form an extracellular and intracellular gate.

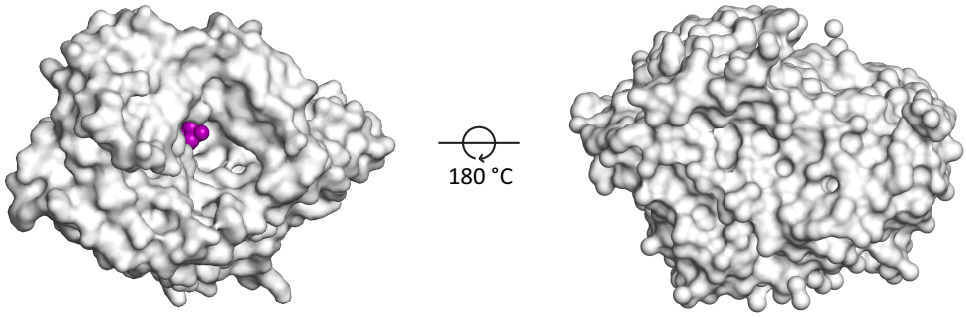


Figure 4.20. A surface representation of the *P. mirabilis* SSS sialic acid transporter. A view from the top of the *P. mirabilis* SSS sialic acid transporter (left) shows sialic acid (purple spheres) binding in a deep cavity that will be open and exposed on the extracellular side. The bottom of the transporter (right) is completely occluded so that the sialic acid will not be transported into the intracellular environment in this outward facing conformation.

Based on the proposed mechanism of sodium and solute symport described in Section 4.1.5, our outward facing and open conformation does not currently fit within this model. Elucidation of this novel conformation (Figure 4.21) suggests that substrate binding does not directly result in closure of the gate and it is mediated in some other way; which may be related to oligomerisation, given that the protein has been demonstrated to have variable oligomeric states.

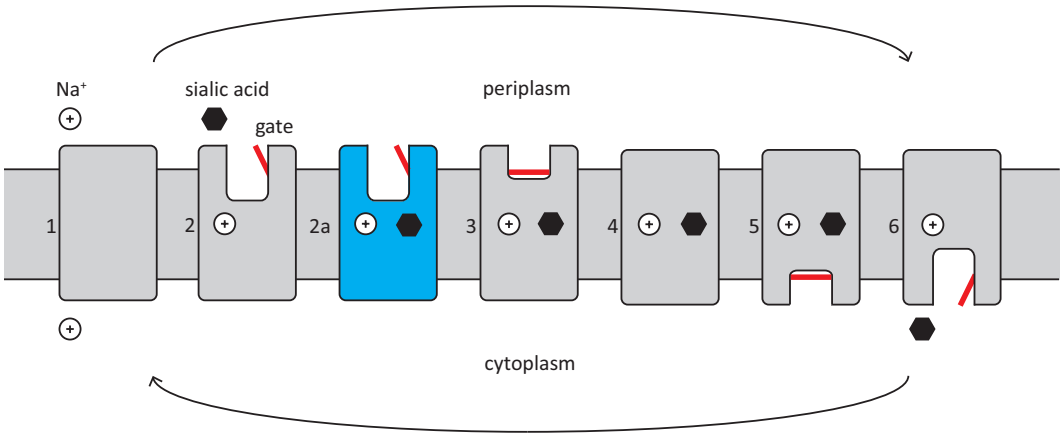


Figure 4.21. A new proposed mechanism of sodium and solute transport. A novel conformation (position 2a) has been discovered in the *P. mirabilis* SSS sialic acid transporter structure. Here, the substrate is bound in the transporter in an outward facing conformation, but the gate remains open.

4.3 Summary

This chapter explored three sialic acid transporter proteins from two distinct gene families, with the aim of gathering structural data to elucidate how sialic acid is transported across the cytoplasmic membrane of bacterial pathogens. Because membrane proteins are notoriously difficult to work with, the *Y. pestis* NanT, *S. aureus* SSS and *P. mirabilis* SSS sialic acid transporters were explored in parallel. Following the development of a successful overexpression, solubilisation and purification protocol for these proteins, the oligomeric structure of these transporter proteins was investigated. The *Y. pestis* NanT sialic acid transporter is shown to be a single species in solution, while the *S. aureus* and *P. mirabilis* SSS sialic acid transporters are shown to form a self-association in solution. In addition, a possible protein-protein interaction was demonstrated for the *S. aureus* SSS sialic acid transporter and MRSA *N*-acetylneuraminate lyase enzyme.

The structure of the *P. mirabilis* SSS sialic acid was solved collaboratively, which is the first known sialic acid transporter structure to be solved. The packing of the crystal lattice demonstrates a possible interface by which oligomerisation may occur. The structure was solved in complex with sialic acid and two sodium ions, providing insight into how this transporter mediates the movement of sialic acid across the membrane. In addition, the structure presents an outward facing and open conformation, a novel conformation among other sodium and solute symporters. These data present a significant step forward in the field of sialobiology and will underpin the design of future antibiotic drugs.

4.4 References

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R. & Iwata, S. (2003). Structure and mechanism of the lactose permease of *Escherichia coli*. *Science, 301*, 610-615.

Abramson, J. & Wright, E. M. (2009). Structure and function of $Na^+/^-$symporters with inverted repeats. *Current opinion in structural biology, 19*, 425-432.

Chae, P. S., Rasmussen, S. G., Rana, R. R., Gotfryd, K., Chandra, R., Goren, M. A., Kruse, A. C., Nurva, S., Loland, C. J., Pierre, Y., Drew, D., Popot, J. L., Picot, D., Fox, B. G., Guan, L., Gether, U., Byrne, B., Kobilka, B. & Gellman, S. H. (2010). Maltose-neopentyl glycol (MNG) amphiphiles for solubilisation, stabilisation and crystallisation of membrane proteins. *Nature Methods, 7*, 1003-1008.

Chang, D. E., Smalley, D. J., Tucker, D. L., Leatham, M. P., Norris, W. E., Stevenson, S. J., Anderson, A. B., Grissom, J. E., Laux, D. C., Cohen, P. S. & Conway, T. (2004). Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 7427-7432.

Drew, D., Lerch, M., Kunji, E., Slotboom, D. J. & de Gier, J. W. (2006). Optimisation of membrane protein overexpression and purification using GFP fusions. *Nature Methods, 3*, 303-313.

Drew, D., von Heijne, G., Nordlund, P. & de Gier, J. W. (2001). Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. *FEBS Letters, 507*, 220-224.

Durek, P. & Walther, D. (2008). The integrated analysis of metabolic and protein interaction networks reveals novel molecular organising principles. *BioMed Central Systems Biology, 2*, 100.

Ebel, C. (2011). Sedimentation velocity to characterise surfactants and solubilised membrane proteins. *Methods, 54*, 56-66.

Eskandari, S., Loo, D. D., Dai, G., Levy, O., Wright, E. M. & Carrasco, N. (1997). Thyroid $Na^+/^-$ symporter. *Journal of Biological Chemistry, 272*, 27230-27238.

Faham, S., Watanabe, A., Besserer, G. M., Cascio, D., Specht, A., Hirayama, B. A., Wright, E. M. & Abramson, J. (2008). The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na^+/sugar symport. *Science, 321*, 810-814.

Fuller, T. E., Kennedy, M. J. & Lowery, D. E. (2000). Identification of *Pasteurella multocida* virulence genes in a septicemic mouse model using signature-tagged mutagenesis. *Microbial Pathogenesis, 29*, 25-38.

Giacalone, M. J., Gentile, A. M., Lovitt, B. T., Berkley, N. L., Gunderson, C. W. & Surber, M. W. (2006). Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *Biotechniques, 40*, 355-364.

Gordon, E., Horsefield, R., Swarts, H. G. P., De Pont, J. J. H. H. M., Neutze, R. & Snijder, A. (2008). Effective high-throughput overproduction of membrane proteins in *Escherichia coli*. *Protein expression and purification, 62*, 1-8.

Higgins, C. F. (1992). ABC Transporters - from microorganisms to man. *Annual Review of Cell Biology, 8*, 67-113.

Hjelm, A., Schlegel, S., Baumgarten, T., Klepsch, M., Wickstrom, D., Drew, D. & de Gier, J. W. (2013). Optimising *E. coli*-based membrane protein production using Lemo21(DE3) and GFP-fusions. *Methods in Molecular Biology, 1033*, 381-400.

Hsieh, J. M., Besserer, G. M., Madej, M. G., Bui, H. Q., Kwon, S. & Abramson, J. (2010). Bridging the gap: a GFP-based strategy for overexpression and purification of membrane proteins with intra and extracellular C-termini. *Protein Science, 19*, 868-880.

Huang, Y., Lemieux, M. J., Song, J., Auer, M. & Wang, D. N. (2003). Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science, 301*, 616-620.

Kawate, T. & Gouaux, E. (2006). Fluorescence-detection size-exclusion chromatography for precrystallisation screening of integral membrane proteins. *Structure, 14*, 673-681.

Kelly, D. J. & Thomas, G. H. (2001). The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *Federation of European Microbiological Societies Microbiology Reviews, 25*, 405-424.

Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Cambridge, The Royal Society of Chemistry.

Le Roy, A., Wang, K., Schaack, B., Schuck, P., Breyton, C. & Ebel, C. (2015). AUC and small-angle scattering for membrane proteins. *Methods in Enzymology, 562*, 257-286.

Lebowitz, J., Lewis, M. S. & Schuck, P. (2002). Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Science, 11*, 2067-2079.

Mackenzie, B., Loo, D. D. & Wright, E. M. (1998). Relationships between Na^+/glucose cotransporter (SGLT1) currents and fluxes. *Journal of Membrane Biology, 162*, 101-106.

Maiden, M. C., Davis, E. O., Baldwin, S. A., Moore, D. C. & Henderson, P. J. (1987). Mammalian and bacterial sugar transport proteins are homologous. *Nature, 325*, 641-643.

Martinez, J., Steenbergen, S. & Vimr, E. (1995). Derived structure of the putative sialic acid transporter from *Escherichia coli* predicts a novel sugar permease domain. *Journal of Bacteriology, 177*, 6005-6010.

Mulligan, C., Fischer, M. & Thomas, G. H. (2011). Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *Federation of European Microbiological Societies Microbiology Reviews, 35*, 68-86.

Mulligan, C., Leech, A. P., Kelly, D. J. & Thomas, G. H. (2012). The membrane proteins SiaQ and SiaM form an essential stoichiometric complex in the sialic acid tripartite ATP-independent periplasmic (TRAP) transporter SiaPQM (VC1777–1779) from *Vibrio cholerae*. *Journal of Biological Chemistry, 287*, 3598-3608.

Olson, M. E., King, J. M., Yahr, T. L. & Horswill, A. R. (2013). Sialic acid catabolism in *Staphylococcus aureus*. *Journal of Bacteriology, 195*, 1779-1788.

Ostermeier, C. & Michel, H. (1997). Crystallisation of membrane proteins. *Current opinion in structural biology, 7*, 697-701.

Prive, G. G. (2007). Detergents for the stabilisation and crystallisation of membrane proteins. *Methods, 41*, 388-397.

Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G. & Deber, C. M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 1760-1765.

Ressl, S., Terwisscha van Scheltinga, A. C., Vonrhein, C., Ott, V. & Ziegler, C. (2009). Molecular basis of transport and regulation in the Na^+/betaine symporter BetP. *Nature, 458*, 47-52.

Ricker, R. D. & Sandoval, L. A. (1996). Fast, reproducible size-exclusion chromatography of biological macromolecules. *Journal of Chromatography A, 743*, 43-50.

Rivas, G., Stafford, W. & Minton, A. P. (1999). Characterisation of heterologous protein-protein interactions using analytical ultracentrifugation. *Methods, 19*, 194-212.

Schlegel, S., Lofblom, J., Lee, C., Hjelm, A., Klepsch, M., Strous, M., Drew, D., Slotboom, D. J. & de Gier, J. W. (2012). Optimising membrane protein overexpression in the *Escherichia coli* strain Lemo21(DE3). *Journal of Molecular Biology, 423*, 648-659.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modelling. *Biophysical Journal, 78*, 1606-1619.

Schultz, S. G. & Curran, P. F. (1970). Coupled transport of sodium and organic solutes. *Physiological Reviews, 50*, 637-718.

Severi, E., Hood, D. W. & Thomas, G. H. (2007). Sialic acid utilisation by bacterial pathogens. *Microbiology, 153*, 2817-2822.

Severi, E., Hosie, A. H. F., Hawkhead, J. A. & Thomas, G. H. (2010). Characterisation of a novel sialic acid transporter of the sodium solute symporter (SSS) family and *in vivo* comparison with known bacterial sialic acid transporters. *Federation of European Microbiological Societies Microbiology Letters, 304*, 47-54.

Severi, E., Randle, G., Kivlin, P., Whitfield, K., Young, R., Moxon, R., Kelly, D., Hood, D. & Thomas, G. H. (2005). Sialic acid transport in *Haemophilus influenzae* is essential for lipopolysaccharide sialylation and serum resistance and is dependent on a novel tripartite ATP-independent periplasmic transporter. *Molecular Microbiology, 58*, 1173-1185.

Slotboom, D. J., Duurkens, R. H., Olieman, K. & Erkens, G. B. (2008). Static light scattering to characterise membrane proteins in detergent solution. *Methods, 46*, 73-82.

Turk, E., Kim, O., le Coutre, J., Whitelegge, J. P., Eskandari, S., Lam, J. T., Kreman, M., Zampighi, G., Faull, K. F. & Wright, E. M. (2000). Molecular characterisation of *Vibrio parahaemolyticus* vSGLT. *Journal of Biological Chemistry, 275*, 25711-25716.

Vimr, E. R., Kalivoda, K. A., Deszo, E. L. & Steenbergen, S. M. (2004). Diversity of microbial sialic acid metabolism. *Microbiology and Molecular Biology Reviews, 68*, 132-153.

Vimr, E. R. & Troy, F. A. (1985). Identification of an inducible catabolic system for sialic acids (*nan*) in *Escherichia coli*. *Journal of Bacteriology, 164*, 845-853.

Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Hogbom, M., van Wijk, K. J., Slotboom, D. J., Persson, J. O. & de Gier, J. W. (2008). Tuning *Escherichia coli* for membrane protein overexpression. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 14371-14376.

Wanner, B. L., Kodaira, R. & Neidhardt, F. C. (1977). Physiological regulation of a decontrolled *lac* operon. *Journal of Bacteriology, 130*, 212-222.

Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E. P., Rutherford, N. G., Hadden, J. M., O'Reilly, J., Ma, P., Saidijam, M., Patching, S. G., Hope, R. J., Norbertczak, H. T., Roach, P. C., Iwata, S., Henderson, P. J. & Cameron, A. D. (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science, 322*, 709-713.

Wiener, M. C. (2004). A pedestrian guide to membrane protein crystallisation. *Methods, 34*, 364-372.

Wright, E. M., Loo, D. D. F., Hirayama, B. A. & Turk, E. (2004). Surprising versatility of Na^+-glucose cotransporters: SLC5. *Physiology, 19*, 370-376.

Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. (2005). Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature, 437*, 215-223.

Chapter Five

Conclusions and future perspectives

5.1 Overview

Using enzymology and structural biology, this thesis explores the import and degradation of sialic acid in clinically important human bacterial pathogens. Bacterial pathogens that reside in heavily sialylated environments have evolved to degrade host-derived sialic acid as a nutrient source (Vimr *et al.*, 2004; Almagro-Moreno & Boyd, 2009). First, this requires sialic acid to be imported into the bacterial cell across the cytoplasmic membrane. Two types of sialic acid transporters were investigated in this thesis, including the sugar cation symporter (NanT) from *Yersinia pestis* and the sodium solute symporters (SSS) from *Staphylococcus aureus* and *Proteus mirabilis* (Figure 5.1 A). Following the import of sialic acid, the catalytic degradation pathway is initiated. The first enzyme [*N*-acetylneuraminate lyase (Figure 5.1 B)] and the third enzyme [*N*-acetylmannosamine-6-phosphate 2-epimerase (Figure 5.1 C)] of this pathway from methicillin-resistant *S. aureus* (MRSA) were also explored. Given the importance of the import and degradation of sialic acid for the colonisation and persistence of bacterial pathogens, surprisingly little structural or functional data exists to explain how bacteria import and degrade sialic acids. This knowledge will inform inhibitor design and improve our understanding of this important, but poorly understood, metabolic pathway.

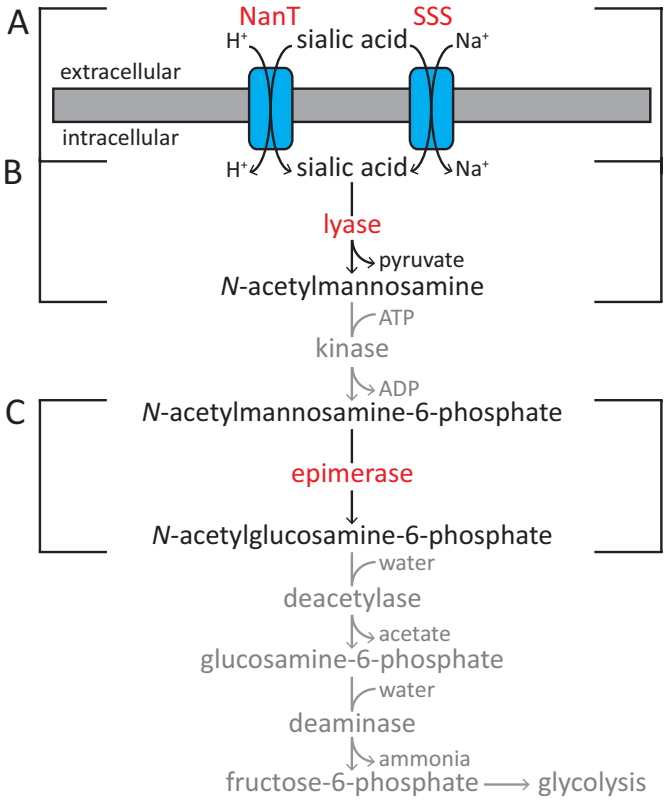


Figure 5.1. The aspects of sialic acid import and degradation investigated in this thesis. (A) Sialic acid import by the NanT and SSS sialic acid transporters. **(B)** *N*-acetylneuraminate lyase is the first enzyme of the sialic acid degradation pathway. **(C)** *N*-acetylmannosamine-6-phosphate 2-epimerase is the third enzyme of the sialic acid degradation pathway.

5.1.1 The first reported structure of a sialic acid transporter

Before sialic acid is degraded as a nutrient source, it must be transported across the cytoplasmic membrane and into the bacterial cell. Sialic acid is a negatively charged molecule that cannot diffuse freely through the membrane bilayer. A dedicated sialic acid transporter protein embedded within the cytoplasmic membrane is responsible for its import (Severi *et al.*, 2007). In gram negative bacterial pathogens, such as *Y. pestis* and *P. mirabilis*, sialic acid must first cross the outer cell membrane into the periplasmic space before it can be imported into the bacterial cell. This process is not well understood, however, a sialic acid inducible porin has recently been implicated in the transport of sialic acid across the outer membrane in *Escherichia coli* (Condemine *et al.*, 2005; Wirth *et al.*,

2009). How this is achieved and the apparent processing of sialic acid in the periplasmic space represents an exciting avenue for future research.

However, the biggest gap in the field is the paucity of structural or functional data that explains how sialic acid transporters mediate the movement of sialic acid across the cytoplasmic membrane, despite a number of groups addressing this issue around the world. The lack of progress is probably because membrane proteins are notoriously challenging to work with (Gordon *et al.*, 2008). To address this, protocols to overexpress, solubilise and purify the *Y. pestis* NanT, *S. aureus* SSS and *P. mirabilis* SSS sialic acid transporters were developed. This is a significant step towards understanding the structure and function of these transporters.

Following purification, solution studies demonstrate the oligomeric structure of these sialic acid transporters for the first time. Size exclusion chromatography and sedimentation velocity experiments using analytical ultracentrifugation demonstrate that the *Y. pestis* NanT sialic acid transporter exists as a single species in solution, whereas the *S. aureus* and *P. mirabilis* SSS sialic acid transporters appear to be in a self-association. Based on the sedimentation velocity data, it is likely that these transporters exist in a monomer-dimer equilibrium, although higher oligomeric structures were also observed for the *S. aureus* SSS sialic acid transporter. Nonetheless, these exciting results suggest that oligomerisation may be important for the SSS sialic acid transporters in nature.

Excitingly, this thesis reports the structure of the *P. mirabilis* SSS sialic acid transporter, which is the first known sialic acid transporter structure to be solved. The structure was solved in complex with sialic acid and two sodium ions, providing insight into how this transporter mediates the movement of sialic acid across the membrane and the stoichiometry of sodium ions required for transport. Importantly, the structure was solved in an outward facing and open conformation, which is a novel conformation among other sodium symporters, that does not currently fit within the proposed mechanism of solute transport. This new conformation suggests that occlusion of the substrate is not a result of substrate binding and it may be mediated in some other way. Given that the SSS sialic acid transporters display variable oligomeric states, perhaps occlusion of the substrate is related to oligomerisation.

The *Y. pestis* NanT sialic acid transporter was successfully crystallised and X-ray diffraction data was collected to 18 Å. However, attempts to crystallise the *S. aureus* SSS sialic acid transporter were unsuccessful. Because the crystallisation of a membrane protein is highly dependent upon the detergent selected (Ostermeier & Michel, 1997; Hsieh *et al.*, 2010), purification and crystallisation of these proteins with alternate detergents should be trialled in the future.

5.1.2 Structure, function and inhibition of MRSA *N*-acetylneuraminate lyase

A detailed structural and functional investigation of MRSA *N*-acetylneuraminate lyase is reported. The kinetic constants reported here for MRSA *N*-acetylneuraminate lyase may be the first for this enzyme, given the ambiguity of a previous kinetic analysis of *S. aureus N*-acetylneuraminate lyase (Timms *et al.*, 2013; Daniels *et al.*, 2014). The values obtained are consistent with the literature for other *N*-acetylneuraminate lyase enzymes. Despite there being a considerable amount of work already published for *N*-acetylneuraminate lyase enzymes, sedimentation velocity experiments demonstrate for the first time that MRSA *N*-acetylneuraminate lyase is tetrameric in solution. This is consistent with other *N*-acetylneuraminate lyase enzymes studied to date and it is also consistent with the crystal structure.

A kinetic and structural analysis of MRSA *N*-acetylneuraminate lyase with a successful first generation inhibitor is presented and the mechanism of inhibition is elucidated for this enzyme. This molecule was demonstrated to be a strong inhibitor for MRSA *N*-acetylneuraminate lyase, which is interesting since it has only been shown to be a moderate inhibitor of *Clostridium perfringens N*-acetylneuraminate lyase, suggesting that there is species-specific inhibition of MRSA *N*-acetylneuraminate lyase. Variable inhibition between *N*-acetylneuraminate lyase enzymes from different organisms may be a result of altered binding modes within respective active sites that is driven by protein dynamics. By comparing the inhibitor bound structure of MRSA *N*-acetylneuraminate lyase with a previously published structure of *Haemophilus influenzae N*-acetylneuraminate lyase, also in complex with this inhibitor (Barbosa *et al.*, 2000), altered binding became apparent.

Currently, the only known regulation for the transporters and enzymes involved in sialic acid degradation is at the transcription level, by a DNA binding protein known as the nanRepressor. Surprisingly, *N*-acetylneuraminate lyase does not appear to be allosterically regulated, despite the related dihydrodipicolinate synthase enzymes being allosterically regulated by a downstream product, lysine (Blickling *et al.*, 1997).

Given that the concentration of free sialic acid in a cell is extremely low (Sillanaukee *et al.*, 1999), the substrate binding affinities determined here and in the literature for *N*-acetylneuraminate lyase enzymes are surprisingly low (3.2-4.1 mM). Thus, I suggest that *N*-acetylneuraminate lyase is regulated at the level of substrate concentration; that is, since the concentration of sialic acid is much lower relative to the Michaelis-Menten constant (K_M), small changes in concentration will have a large effect on flux.

My results suggest that MRSA *N*-acetylneuraminate lyase and the *S. aureus* SSS sialic acid transporter interact at the cytoplasmic membrane. It has long been established that proteins involved in the same metabolic pathway and more importantly the successive enzymes of the pathway, are often involved in protein-protein interactions with one another by forming either transient or permanent multi-enzyme complexes (Durek & Walther, 2008). Sedimentation velocity experiments provided evidence for the formation of a protein-protein interaction between the *S. aureus* SSS sialic acid transporter and MRSA *N*-acetylneuraminate lyase. If a protein-protein interaction does occur between a sialic acid transporter and *N*-acetylneuraminate lyase, the low substrate binding affinity may be an emergent property of this interaction.

5.1.3 A novel mechanism for carbohydrate epimerase enzymes

Currently, there is very little structural and functional information available for *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes. To address this, the structure and catalytic mechanism of *N*-aetylmannosamine-6-phosphate 2-epimerase from MRSA was investigated using structural biology and enzymology. A multi-enzyme coupled assay was developed that reports *N*-acetylmannosamine-6-phosphate 2-epimerase activity. Using this assay, the first kinetic constants were reported for MRSA *N*-acetylmannosamine-6-

phosphate 2-epimerase. This enzyme was shown to have a considerably higher binding affinity for its substrate than the *C. perfringens N*-acetylmannosamine-6-phosphate 2-epimerase, which was kinetically analysed by proton nuclear magnetic resonance spectroscopy (^{1}H NMR). These types of assays are both time consuming and resource intensive (Exnowitz *et al.*, 2012), whereas the multi-enzyme coupled assay presents a convenient method to assay activity in real time, which will be suitable for testing inhibitors in the future.

For the first time, solution studies demonstrated that MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase is dimeric, which was confirmed by the crystal structure. Using the structure of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase in combination with sequence analysis, residues that may be important for catalysis were probed and their involvement in catalysis was assessed by mutagenesis and kinetic analysis. During the course of this thesis, it was proposed that *N*-acetylmannosamine-6-phosphate 2-epimerase enzymes catalyse their chemistry *via* a deprotonation/reprotonation mechanism (Pelissier *et al.*, 2014). However, there was not sufficient experimental evidence to prove this notion.

Instead, I propose that catalysis occurs *via* a proton displacement mechanism mediated by the substrate. Importantly, a role for the residues that were deemed catalytically essential by kinetic analysis can be predicted if catalysis does occur in this way. This type of mechanism appears to be a novel strategy for carbohydrate epimerase enzymes. The usual mechanisms employed by such epimerases include deprotonation/reprotonation, formation of a transient keto intermediate, carbon-carbon bond cleavage, nucleotide elimination and mutarotation (Allard *et al.*, 2001; Samuel & Tanner, 2002). To confirm this hypothesis, crystallographic analyses of the wild type and mutant enzymes with substrate and derivatives will be attempted. This knowledge could be used to inform inhibitor design.

5.2 References

Allard, S. T., Giraud, M. F. & Naismith, J. H. (2001). Epimerases: structure, function and mechanism. *Cellular and Molecular Life Sciences, 58*, 1650-1665.

Almagro-Moreno, S. & Boyd, E. F. (2009). Insights into the evolution of sialic acid catabolism among bacteria. *BioMed Central Evolutionary Biology, 9*, 118-133.

Barbosa, J., Smith, B., DeGori, R., Ooi, H., Marcuccio, S., Campi, E., Jackson, W., Brossmer, R., Sommer, M. & Lawrence, M. (2000). Active site modulation in the *N*-acetylneuraminate lyase sub-family as revealed by the structure of the inhibitor-complexed *Haemophilus influenzae* enzyme. *Journal of Molecular Biology, 303*, 405-421.

Blickling, S., Renner, C., Laber, B., Pohlenz, H. D., Holak, T. A. & Huber, R. (1997). Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochemistry, 36*, 24-33.

Condemine, G., Berrier, C., Plumbridge, J. & Ghazi, A. (2005). Function and expression of an *N*-acetylneuraminic acid-inducible outer membrane channel in *Escherichia coli*. *Journal of Bacteriology, 187*, 1959-1965.

Daniels, A. D., Campeotto, I., van der Kamp, M. W., Bolt, A. H., Trinh, C. H., Phillips, S. E., Pearson, A. R., Nelson, A., Mulholland, A. J. & Berry, A. (2014). Reaction mechanism of *N*-acetylneuraminic acid lyase revealed by a combination of crystallography, QM/MM simulation, and mutagenesis. *American Chemical Society Chemical Biology, 9*, 1025-1032.

Durek, P. & Walther, D. (2008). The integrated analysis of metabolic and protein interaction networks reveals novel molecular organising principles. *BioMed Central Systems Biology, 2*, 100.

Exnowitz, F., Meyer, B. & Hackl, T. (2012). NMR for direct determination of K_m and V_{max} of enzyme reactions based on the Lambert W function-analysis of progress curves. *Biochimica et Biophysica Acta, 1824*, 443-449.

Gordon, E., Horsefield, R., Swarts, H. G. P., De Pont, J. J. H. H. M., Neutze, R. & Snijder, A. (2008). Effective high-throughput overproduction of membrane proteins in *Escherichia coli*. *Protein expression and purification, 62*, 1-8.

Hsieh, J. M., Besserer, G. M., Madej, M. G., Bui, H. Q., Kwon, S. & Abramson, J. (2010). Bridging the gap: a GFP-based strategy for overexpression and purification of

membrane proteins with intra and extracellular C-termini. *Protein Science, 19*, 868-880.

Ostermeier, C. & Michel, H. (1997). Crystallisation of membrane proteins. *Current opinion in structural biology, 7*, 697-701.

Pelissier, M. C., Sebban-Kreuzer, C., Guerlesquin, F., Brannigan, J. A., Bourne, Y. & Vincent, F. (2014). Structural and functional characterisation of the *Clostridium perfringens N*-acetylmannosamine-6-phosphate 2-epimerase essential for the sialic acid salvage pathway. *Journal of Biological Chemistry, 289*, 35215-35224.

Samuel, J. & Tanner, M. E. (2002). Mechanistic aspects of enzymatic carbohydrate epimerisation. *Nature Product Reports, 19*, 261-277.

Severi, E., Hood, D. W. & Thomas, G. H. (2007). Sialic acid utilisation by bacterial pathogens. *Microbiology, 153*, 2817-2822.

Sillanaukee, P., Ponnio, M. & Jaaskelainen, I. P. (1999). Occurrence of sialic acids in healthy humans and different disorders. *European Journal of Clinical Investigation, 29*, 413-425.

Timms, N., Windle, C. L., Polyakova, A., Ault, J. R., Trinh, C. H., Pearson, A. R., Nelson, A. & Berry, A. (2013). Structural insights into the recovery of aldolase activity in *N*-acetylneuraminic acid lyase by replacement of the catalytically active lysine with gamma-thialysine by using a chemical mutagenesis strategy. *Chembiochem: a European Journal of Chemical Biology, 14*, 474-481.

Vimr, E. R., Kalivoda, K. A., Deszo, E. L. & Steenbergen, S. M. (2004). Diversity of microbial sialic acid metabolism. *Microbiology and Molecular Biology Reviews, 68*, 132-153.

Wirth, C., Condemine, G., Boiteux, C., Berneche, S., Schirmer, T. & Peneff, C. M. (2009). NanC crystal structure, a model for outer-membrane channels of the acidic sugar-specific KdgM porin family. *Journal of Molecular Biology, 394*, 718-731.

Chapter Six

Experimental procedures

6.1 Experimental reagents

6.1.1 Chemical reagents

Sialic acid and *N*-acetylmannosamine-6-phosphate were purchased from Carbosynth. Chemicals used for media preparation were sourced from Invitrogen. All antibiotics were purchased from Sigma-Aldrich. Roche supplied cOmplete™, ethylenediaminetetraacetic acid (EDTA) free protease inhibitor tablets.

6.1.2 Biological reagents

Plasmids containing genes of interest were purchased from GenScript or GeneArt and oligonucleotides were ordered from Integrated DNA Technologies. Chemically competent *Escherichia coli* XL-1 Blue cells were obtained from Agilent Technologies. Chemically competent *E. coli* BL21(DE3) and Lemo21(DE3) cells were obtained from New England Biolabs. All restriction enzymes and Phusion® high fidelity deoxyribonucleic acid (DNA) polymerase were also obtained from New England Biolabs. TOPO TA® cloning kits and the PureLink® quick plasmid miniprep kits were purchased from Invitrogen. Agarose gel DNA extraction kits were supplied by Roche. Ligase was purchased from Takara. DNase I, Bovine serum albumin, glucose-6-phosphate dehydrogenase, lactate dehydrogenase and phosphoglucoisomerase were obtained from Sigma-Aldrich. Equipment required for sodium dodecyl sulfate polyacrylamide gel electrophoresis, including sample buffers, running buffers, precast gels and protein molecular weight markers were supplied by Thermofisher Scientific. Bioline supplied DNA markers and sample loading buffer.

SimplyBlue™ SafeStain was purchased from Invitrogen. SYPRO® Orange and SYBR® Safe DNA gel stain were also purchased from Invitrogen. Bradford dye reagent was supplied by BioRad.

6.1.3 General materials

All prepacked chromatography columns were purchased in various sizes from GE Healthcare. Nickel nitrilotriacetic acid (Ni-NTA) agarose resin and empty Poly-Prep® columns for gravity flow chromatography were obtained from Qiagen and BioRad, respectively. BioRad also supplied 96 well thick wall plates, while Sigma-Aldrich supplied Corning® 96 well clear bottom plates. Crystallisation screens that were used in-house and solutions used for cryo protection were purchased from Molecular Dimensions. Hampton Research supplied crystallisation plates and cryo loops. All solutions were made with Millipore purified water, which is referred to as MilliQ water throughout this thesis. Solutions that required filtering were passed through 0.22 μm syringe filters obtained from Millipore. Proteins were concentrated using appropriate molecular weight cut off spin concentrators obtained from Sartorius.

6.2 General methods

6.2.1 Bioinformatic analyses of nucleotide and amino acid sequences

A variety of bioinformatic tools were employed throughout this study. For the determination of nucleotide and amino acid sequences, the basic local alignment search tool (*BLAST*) programs '*BLASTn*' and '*BLASTp*' were utilised respectively (Karlin & Altschul, 1990; Karlin & Altschul, 1993) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple sequence alignments to investigate the similarity between nucleotide and amino acid sequences were conducted using the programs *MUSCLE* (Edgar, 2004), or *ClustalW* (Larkin *et al.*, 2007) (http:// www.genome.jp/tools/clustalw/), with manual editing.

6.2.2 Centrifugation

All centrifugation was carried out using a Sorvall™ RC 6 Plus Centrifuge (Thermo Scientific), Eppendorf™ Centrifuge 5810 R (Thermo Fisher Scientific) or an Optima™ L-90K Ultracentrifuge (Beckman Coulter).

6.2.3 Measurement of molecular weight

The masses of purified proteins were analysed on a MaXis 3G quadrople time of flight mass spectrometer, equipped with an electrospray ionisation source (Bruker Daltonics). Proteins to be measured were diluted to approximately 1 mg/mL with MilliQ water.

6.2.4 pH measurement

The pH of solutions was measured at room temperature using an Ultrabasic Benchtop pH meter (Denver Instrument). The pH was adjusted using 1 M or 10 M hydrochloric acid if the pH of the solution needed to be decreased, or 1 M or 10 M sodium hydroxide if the pH of the solution needed to be increased.

6.2.5 Sterilisation of media and equipment

Media for bacterial cultures and the equipment to be used with it was autoclaved at 121 °C for 15 minutes. Solutions were prepared with MilliQ water. If the solution could not be autoclaved, it was filtered through a 0.22 μm syringe filter. Standard aseptic technique was used for each experiment and controls were used to monitor for any possible contamination.

6.3 Microbiology

6.3.1 Bacterial strains

Five bacterial strains of *E. coli* were routinely used in the research conducted for this thesis; these are described in Table 2.1.

Table 2.1 Bacterial strains used throughout this thesis. The strain name and associated genotype is presented.

E. coli strain	Genotype
BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS* *λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*
Lemo21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/pLemo*(CamR) *λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5* pLemo = pACYC184-*PrhaBAD-lysY*
XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI*^{q}*ZΔM15* Tn10 (Tet^{r})]

6.3.2 Antibiotics

Antibiotic stock solutions were prepared at a 1 000 × concentration and sterilised through a 0.22 µm syringe filter before storage at -20 °C. Antibiotics used for bacterial selection throughout this thesis are listed in Table 2.2.

Table 2.2 Antibiotics used for bacterial selection. The stock concentration, working concentration and solvent used to dissolve the respective antibiotic are listed.

Antibiotic	Stock Concentration (mg/mL)	Working Concentration (µg/mL)	Solvent
ampicillin	100	100	MilliQ water
kanamycin	50	50	MilliQ water
chloramphenicol	34	34	ethanol

6.3.3 Media

6.3.3.1 Luria bertani medium

Luria bertani (LB) base was purchased in a ready to use powdered form. For every 1 L of LB medium to be prepared, 20 g of LB base was added to MilliQ water. This was sterilised by autoclaving and then stored at room temperature.

6.3.3.2 LB agar

LB base and agar were mixed and prepared in MilliQ water. For every 1 L of LB agar prepared, 20 g of LB base, and 12 g of agar were added to MilliQ water. This was sterilised by autoclaving. Molten media was supplemented with the appropriate antibiotic, mixed by gentle swirling and poured into sterile petri dishes in either a laminar flow hood, or close to a flame. Plates were left to set in sterile conditions before being sealed with parafilm and stored at 4 °C for up to two weeks.

6.3.3.3 Terrific broth medium

Terrific broth (TB) medium was prepared by the addition of 12 g tryptone, 24 g yeast extract, 2.2 g monopotassium phosphate, 9.4 g dipotassium phosphate, and 4 mL glycerol to every 1 L of MilliQ water. The pH was adjusted to 7.2 with 10 M sodium hydroxide. TB medium was sterilised by autoclaving prior to being stored at room temperature.

6.3.3.4 Super optimal broth with catabolite repression medium

Super optimal broth with catabolite repression (SOC) medium was prepared by the addition of 20 g tryptone, 5 g yeast extract, 0.5 g sodium chloride, 2.5 mL 1 M potassium chloride, 10 mL 1 M magnesium chloride, 10 mL 1 M magnesium sulfate, and 20 mL glucose to every 1 L of MilliQ water. The pH was adjusted to 7.0 with 10 M sodium hydroxide. SOC medium was sterilised by autoclaving and then aliquoted into 1.7 mL eppendorf tubes in either a laminar flow hood, or close to a flame. SOC medium was stored at -20 °C.

6.3.4 Transformation of chemically competent bacterial cells

All strains of chemically competent *E. coli* used in this thesis were commercially prepared, supplied and transformed according to manufacturer's instructions. Briefly, competent cells were thawed on ice immediately prior to transformation. Approximately 100 ng of DNA was added to an aliquot of competent cells and incubated on ice for 20 minutes prior to a heat shock at 42 °C for 45 seconds. Cells were returned to ice for 5 minutes to cool. Subsequently, 400 µL of SOC media was added, and the cells were then incubated at 37 °C and 200 revolutions per minute (rpm) for 1 hour. Aliquots were plated onto antibiotic selective LB agar plates and incubated at 37 °C overnight.

6.3.5 Bacterial culturing

Starter cultures were prepared by inoculating a single bacterial colony in 5 mL of desired media with the appropriate antibiotic. These were incubated for 8 hours at 37 °C and 200 rpm. Overnight cultures containing the appropriate media and antibiotic were inoculated with a 1/50 dilution of the starter culture. These were incubated at either 37 °C for soluble proteins, or 30 °C for membrane proteins, and 200 rpm overnight. Large scale cultures containing the appropriate media and antibiotic were further inoculated with a 1/50 dilution of the overnight culture. These were incubated at 37 °C and 200 rpm until an appropriate optical density at 600 nm (OD_{600}) was reached. Once reached, protein expression was induced and the cultures were incubated at a temperature appropriate for the respective construct and at 200 rpm overnight (Section 6.6.1).

6.3.6 Preparation of glycerol stocks

Bacterial strains were stored by harvesting an aliquot of overnight culture in a sterile 1.7 mL eppendorf tube. Glycerol was added and combined to a final concentration of 15% (v/v). The tubes were snap frozen in liquid nitrogen and stored at -80 °C.

6.4 Molecular biology

6.4.1 Plasmids

The *nanA* gene encoding *N*-acetylneuraminate lyase from methicillin-resistant *Staphylococcus aureus* (MRSA) MRSA252 was previously amplified from genomic DNA and cloned into a pET11a expression vector. This construct was kindly supplied by Professor Matthew Perugini (University of Melbourne, Australia). The *nanE*, *nagA* and *nagB* genes respectively encoding *N*-acetylmannosamine-6-phosphate 2-epimerase, *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase from MRSA USA300 were purchased from GenScript in the pUC57-Kan cloning vector and subsequently cloned into the pET30ΔSE expression vector (Suzuki *et al.*, 2014). A more detailed description of this cloning is provided in Chapter Three, Section 3.2.1. Mutated variants of the *nanE* gene coding for single amino acid substitutions in *N*-acetylmannosamine-6-phosphate 2-epimerase were also purchased from GenScript. These genes were supplied in the pET30a(+) expression vector.

The gene encoding the sugar proton symporter (NanT) responsible for the transport of sialic acid from *Yersinia pestis* was purchased from GeneArt in the pMA-RQ cloning vector. This gene was then sub-cloned into the pWarf(−) expression vector (Hsieh *et al.*, 2010). A more detailed description of this cloning is provided in Chapter Four, Section 4.2.1. The genes encoding the sodium solute symporters (SSS) responsible for the transport of sialic acid from *S. aureus* and *Proteus mirabilis* were previously sub-cloned into the pWarf(−) expression vector and kindly supplied by Dr Rosmarie Friemann (University of Gothenburg, Sweden).

6.4.2 Plasmid extraction and isolation

Plasmids were purified from 5 mL starter cultures of XL-1 Blue *E. coli* using a PureLink® Quick Plasmid Miniprep Kit. The concentration of purified plasmid was measured by absorption at 260 nm using a NanoDrop-1 000 spectrophotometer. The plasmid DNA was sequence verified by dideoxynucleotide sequencing (Section 6.4.3).

6.4.3 Dideoxynucleotide sequencing

Canterbury Sequencing and Genotyping at the University of Canterbury provided dideoxynucleotide sequencing. Approximately 300 ng of purified double stranded plasmid and 3.2 μM of the relevant primer were supplied for each sequencing reaction to be carried out by a 3130xl Genetic Analyser and BigDye Terminator v3.1 (Applied Biosystems).

6.4.4 Primers

Oligonucleotide primer sequences were designed so that the T_M of each primer pair were within 2 °C of each other, the GC content was higher than 40% and the formation of secondary structure elements could be avoided. Primers were adjusted to 100 μM per tube by resuspending dessicated pellets in sterile MilliQ water and then diluted to a working stock of 10 μM with sterile MilliQ water. Primer solutions were stored at -20 °C.

6.4.5 Polymerase chain reaction

The polymerase chain reaction (PCR) was performed using Platinum® *taq* high fidelity DNA polymerase, according to the supplied manual. Briefly, the reaction mixture was composed of 1 × high fidelity PCR buffer, 0.2 mM deoxynucleotide triphosphates (dNTP's), 2 mM magnesium sulfate, 0.2 μM forward primer, 0.2 μM reverse primer, 200 ng template DNA and 1 U of Platinum® *taq* high fidelity DNA polymerase made up to a total volume of 50 μL with sterile MilliQ water. PCR was performed according to the cycling conditions in Table 2.3. The thermal melting temperature (T_M) reflects the lower temperature of the two PCR primers. Resulting PCR products were sub-cloned according to manufacturer's instructions into pCR®2.1-TOPO using a TOPO® TA cloning kit and further sub-cloned into an expression vector (Sections 6.4.6 – 6.4.8).

Table 2.3 PCR cycling reaction. The temperature, time, and number of cycles used in a PCR cycling reaction.

Temperature (°C)	Time	Cycles
94	2 minutes	1
94	15 seconds	
$T_M - 5$	30 seconds	30
68	1 minute per kilobase	
4	as required	1

6.4.6 Restriction digestion

To obtain either a gene of interest, or an empty plasmid for subsequent cloning, restriction digests were performed to cut plasmids at particular recognition sites. Restriction digests were performed according to the manufacturer's instructions. Components of a typical restriction digestion are listed in Table 2.4. Reactions were incubated at 37 °C for 3 hours and then analysed by agarose gel electrophoresis (Section 6.4.7).

Table 2.4 Restriction enzyme digest reaction. Components are listed for a reaction volume of 50 μL.

Reagent	Volume (μL)
10 × reaction buffer	5
plasmid DNA	~2 000 ng DNA
restriction enzyme 1	1
restriction enzyme 2	1
MilliQ Water	total volume 50

6.4.7 Agarose gel electrophoresis for DNA preparation

Following PCR and/or restriction digestion, the DNA fragments were mapped on a 1% (w/v) agarose gel (Section 6.5.1). The gel was loaded with digested DNA in loading buffer alongside a DNA ladder, and then electrophoresis was conducted for 40 minutes at 120 volts. DNA fragments were visualised on an ultraviolet (UV) light box and either the digested gene of interest, or digested and empty plasmid, was carefully excised from the gel. DNA was purified from the gel using an agarose gel DNA extraction kit. The concentration of purified DNA was measured by absorption at 260 nm using a NanoDrop-1 000 spectrophotometer.

6.4.8 Ligation reactions

Following purification of the relevant plasmid and gene of interest, a ligation reaction was carried out. Components of a typical ligation reaction are listed in Table 2.5. Reactions were incubated at room temperature for 50 minutes and then transformed into XL-1 Blue competent cells. Positive transformants were identified *via* antibiotic resistance prior to plasmid purification and sequence verification using dideoxynucleotide sequencing. Following verification, the resulting recombinant plasmid was transformed into an appropriate cell line of *E. coli* for protein expression.

Table 2.5 **Ligation reaction.** Components for a reaction volume of 5 μL.

Reagent	Volume (μL)
plasmid DNA	~100 ng
insert DNA	~200 ng
ligase	2
MilliQ Water	total volume 5

6.5 Electrophoresis

6.5.1 Agarose gel electrophoresis

DNA fragments were separated using agarose gel electrophoresis. Agarose gels were prepared by dissolving 1% (w/v) agarose in 1 × 2-Amino-2-hydroxymethyl-propane-1,3-diol (Tris)-acetate-EDTA (TAE) buffer (40 mM Tris, 20 mM glacial acetic acid and 1 mM EDTA). Once the solution reached approximately 50 °C, SYBR® Safe DNA Gel Stain was added to a final concentration of 1 × and the gel was poured and left to set. DNA samples were combined with 6 × DNA gel loading dye to a final concentration of 1 × and loaded into the gel alongside HyperLadder™ 1 kb. Gels were electrophoresed in 1 × TAE buffer at 120 volts for 40 minutes. Bands were visualised by UV transillumination.

6.5.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) to separate proteins was performed using either Novex® 4-12% tris glycine 1.0 mM or Novex® 4-12% bis tris 1.0 mM precast protein gels with 1 × Novex® tris glycine SDS or 1 × Novex® 2-(N-morpholino)ethanesulfonic acid SDS buffer, respectively. Electrophoresis was performed at room temperature using a XCell SureLock® apparatus for 35 minutes at 200 volts for bis tris gels, or 1.5 hours and 125 volts for tris glycine gels. Samples to be electrophoresed on a bis tris gel were resuspended in NuPAGE® lithium dodecyl sulfate sample buffer and NuPAGE® sample reducing agent. Samples to be electrophoresed on a tris glycine gel were resuspended in 1 × Novex ® tris glycine SDS sample buffer. Samples were electrophoresed alongside Novex® sharp prestained protein standards. If in-gel fluorescence needed to be assessed (Drew *et al.*, 2006; Hjelm *et al.*, 2013), then samples were electrophoresed alongside a BenchMark™ fluorescent protein standard. Following electrophoresis, gels were stained using SimplyBlue™ SafeStain. Images of the gels were captured using a CHEMI GENIUS^2 bio-imaging system (Syngene). Chemiluminescence imaging was used to detect in-gel fluorescence.

6.6 Protein biochemistry

6.6.1 Recombinant protein expression

Cultures for protein expression were routinely grown to an appropriate OD_{600} of 0.4 – 0.6 for LB medium and 1.7 – 1.9 for TB medium. Once these OD_{600} levels were achieved, cultures were cooled to an appropriate temperature for expression of the relevant construct. Induction was achieved by the addition of isopropyl-β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM for the expression of membrane proteins, or 1 mM for the expression of soluble proteins.

6.6.2 Harvesting of cells

Following protein expression, large scale cultures were harvested by centrifugation at 6250 g, for 20 minutes, at 4 °C. Cell pellets were washed in ice cold 1 × phosphate buffered saline (PBS) solution and the centrifugation step was repeated. Cell pellets were then snap frozen in liquid nitrogen and stored at -80 °C.

6.6.3 Cell lysis

All cell lysis and purification was performed either on ice or at 4 °C. Cell pellets were thawed and resuspended in approximately 3 mL/g of cell pellet in 1 × PBS solution, or the first buffer to be used in purification. Proteolytic activity was inhibited by the addition of 1 cOmplete™, EDTA free protease inhibitor tablet per 50 mL of cell resuspension. In addition, lysozyme was added to the cell resuspension at a final concentration of 0.5 mg/mL and incubated on ice for 30 minutes. Subsequently, 5 µg/mL of DNaseI and 2 mM magnesium chloride were added to the cell resuspension and incubated on ice for a further 15 minutes. Cells were lysed by sonication on ice using a UP200S Ultrasonic Processor (Hielscher) at 70% amplitude, 0.5 seconds on, 0.5 seconds off, for 10 minutes. Any unbroken cells and cell debris were removed by centrifugation at 24 000 g, for 25 minutes, at 4 °C. Supernatant was collected and stored at 4 °C.

6.6.4 Whole cell fluorescence

Following protein expression trials, the OD_{600} of the cell culture was measured and 5 mL of culture was harvested. Cell pellets were resuspended in PBS buffer and 200 µL was transferred to a Corning® 96 well, clear bottom plate for the measurement of fluorescence from a green fluorescent protein (GFP) fusion tag with excitation at λ = 485 nm and emission at λ = 512 nm (Drew *et al.*, 2001; Drew *et al.*, 2006; Hjelm *et al.*, 2013). All fluorescence data was processed as described in Hsieh *et al.* (2010). The fluorescent values (F) were normalised (F_N) against the OD_{600} measurement (Equation 6.1). Furthermore, the fluorescence values were reported as relative fluorescent units (RFU); where the normalised fluorescence (F_{N1}) is divided by the normalised fluorescence of non-induced cells (F_{N2}) that contain the same plasmid (Equation 6.2).

Equation 6.1. $$F_N = \frac{F}{OD_{600}}$$

Equation 6.2. $$RFU = \frac{F_{N1}}{F_{N2}}$$

6.6.5 Harvesting of membranes

For the membrane proteins expressed and studied in this thesis, the membrane fraction was harvested following cell lysis. The supernatant collected in Section 6.6.3 was centrifuged a second time at 24 000 g, for 25 minutes, at 4 °C to clear the supernatant from any residual cells and cell debris. Membranes were separated from the supernatant by ultracentrifugation at 200 000 g, for 1.5 hours, at 4 °C, using a Ti45 rotor. The supernatant was discarded and the membranes were washed in membrane resuspension buffer (1 × PBS, 5 mM dithiothreitol, 6% glycerol and 0.1 µM phenylmethylsulfonyl fluoride). Ultracentrifugation was repeated at 200 000 g, for 1 hour, at 4 °C. The membrane pellet was snap frozen in liquid nitrogen and stored at -80 °C.

6.6.6 Membrane protein solubilisation

The membrane pellet was slowly thawed on ice and 5 mL of membrane resuspension buffer was added for every 1 g of membranes. Once thawed, an appropriate detergent (Section 6.6.7) was added to a final concentration of 2% (w/v). Solubilisation was carried out for 2 hours, at 4 °C, with gentle agitation. The membranes that had not been solubilised were pelleted by centrifugation at 24 000 g, for 1 hour, at 4 °C. The supernatant was carefully separated from the pellet and kept at 4 °C.

6.6.7 Detergent screening

Fluorescence detection size exclusion chromatography (FSEC) was used as an efficient method to identify ideal detergents. Approximately 100 mg of membrane pellets were solubilised in a range of detergents. Following centrifugation, 250 µL of the supernatant was immediately applied to a Superose 6 10/150 column that was pre-equilibrated in membrane resuspension buffer containing the respective detergent at a concentration above the critical micelle concentration. Because FSEC can exploit the unique fluorescence signal of GFP, the eluate was monitored by GFP emission at 512 nm (Drew *et al.*, 2006; Hsieh *et al.*, 2010). FSEC was repeated after 48 hours with a further 250 µL of the supernatant that was kept at 4 °C.

6.6.8 Protein purification

Various chromatographic techniques were performed in this work using ÄKTA protein purification systems and gravity flow purification methods. All protein purification procedures were conducted at 4 °C. Prior to any procedure, the columns as well as the resin used for gravity flow techniques were pre-equilibrated with at least three column volumes of the initial purification buffer. Depending on the type of protein being purified and the purification step being utilised, the cell free supernatant, solubilised membrane fraction, or partially purified protein was loaded on to the appropriate column. To remove any unbound protein, the column was then washed with an additional three column volumes of the initial purification buffer until stable baselines at 280 nm were achieved.

For immobilised metal affinity chromatography (IMAC), anion exchange chromatography and hydrophobic interaction chromatography, the protein was eluted with either an increasing or a step wise gradient of a second comparable buffer containing a much higher concentration of imidazole, or sodium chloride, or a reduced concentration of ammonium sulfate, respectively. The eluate was fractionated and collected in 96 well plates. Following these procedures, fractions were conservatively selected and pooled with respect to their purity, as determined by SDS-PAGE. After the final purification step, all proteins were aliquoted into ice cold 1.7 mL eppendorf tubes and snap frozen in liquid nitrogen for long term storage at -80 °C.

6.6.8.1 Anion exchange chromatography

Anion exchange chromatography was employed as the first purification step for the purification of MRSA *N*-acetylneuraminate lyase, *N*-acetylmannosamine-6-phosphate 2-epimerase, *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase. The components of each buffer used for anion exchange chromatography are described in Table 2.6. The cell free supernatant was applied to a 20 mL anion exchange column containing positively charged Q sepharose resin, which had been pre-equilibrated with anion exchange buffer A. Once the cell free supernatant was loaded onto the column, it was washed with three column volumes of anion exchange buffer A. Bound proteins were eluted with a continuous gradient of anion exchange buffer B, up to 100% over three column volumes. Eluate was collected in fractions and peaks in the 280 nm absorbance trace were analysed for the protein of interest by SDS-PAGE before being pooled.

Table 2.6. Anion exchange chromatography buffers. Components of each buffer are listed.

Buffer	Components
anion exchange buffer A	20 mM Tris, pH 8.0
anion exchange buffer B	20 mM Tris, pH 8.0; 1 M sodium chloride

6.6.8.2 Hydrophobic interaction chromatography

Hydrophobic interaction chromatography was implemented as a second purification step for any proteins that were subjected to anion exchange chromatography as an initial purification step. This includes MRSA *N*-acetylneuraminate lyase, *N*-acetylmannosamine-

6-phosphate 2-epimerase, *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase. The components of each buffer used for hydrophobic interaction chromatography are described in Table 2.7. To a final concentration of 1 M, ammonium sulfate was added to the partially purified protein pooled from the anion exchange chromatography step. The protein was then applied to a phenyl sepharose 20 mL column pre-equilibrated with hydrophobic interaction buffer A, and washed with a further three column volumes of this buffer. Bound proteins were eluted with a continuous gradient of hydrophobic interaction buffer B, up to 100% over three column volumes. Fractions corresponding to any 280 nm absorbance peaks were analysed by SDS-PAGE for the protein of interest prior to being pooled. All proteins purified by anion exchange and hydrophobic interaction chromatographies were next subjected to size exclusion chromatography as the final purification step (Section 6.6.8.7).

Table 2.7. Hydrophobic interaction chromatography buffers. Components of each buffer are listed.

Buffer	Components
hydrophobic interaction buffer A	20 mM Tris, pH 8.0; 1 M ammonium sulfate
hydrophobic interaction buffer B	20 mM Tris, pH 8.0

6.6.8.3 Immobilised metal affinity chromatography

IMAC was employed as the first purification step for the sialic acid transporter proteins studied in this thesis (*Y. pestis* NanT, *P. mirabilis* SSS, and *S. aureus* SSS). This type of chromatography was achieved using a His Trap 5 mL column, pre-equilibrated with Buffer A for each selected protein (Table 2.8). The solubilised membrane fraction was applied to the column. The column was then washed with a further five column volumes of the initial buffer used for pre-equilibration. Depending on the protein, elution from the column was conducted using either an increasing or step wise gradient of an elution buffer B made up of the same components as the initial buffer, but with a higher concentration of imidazole (Table 2.10). For the *P. mirabilis* SSS sialic acid transporter, the column was washed for ten column volumes with 5% of the *P. mirabilis* SSS buffer B before a continuous gradient of elution buffer up to 100% was initiated over twenty column volumes. For the *Y. pestis* NanT and *S. aureus* SSS sialic acid transporters, elution was conducted *via* a step wise gradient using the *Y. pestis* NanT and *S. aureus* SSS buffer B, respectively. This involved washing the column with 10% of the relevant elution buffer over ten column volumes, then

increasing the concentration of elution buffer to 50% over twenty column volumes. Fractions corresponding to any 280 nm absorbance peaks were analysed by SDS-PAGE for the protein of interest prior to being pooled.

Table 2.8. IMAC buffers. Buffer components are listed for each protein subjected to IMAC throughout this thesis. Buffer A refers to the initial buffer used for pre-equilibration of the column. Buffer B refers to the elution buffer for each respective protein.

Buffer	Components
Y. pestis NanT buffer A	70 mM Tris, pH 8.0; 150 mM sodium chloride; 20 mM imidazole; 6% glycerol; 5 mM β-mercaptoethanol; 0.0174% n-dodecyl-β-D-maltopyranoside
Y. pestis NanT buffer B	70 mM Tris, pH 8.0; 150 mM sodium chloride; 500 mM Imidazole; 6% Glycerol; 5 mM β-mercaptoethanol; 0.0174% n-dodecyl-β-D-maltopyranoside
P. mirabilis SSS buffer A	70 mM Tris, pH 8.0; 150 mM sodium chloride; 20 mM imidazole; 6% glycerol; 5 mM β-mercaptoethanol; 0.0174% n-dodecyl-β-D-maltopyranoside
P. mirabilis SSS buffer B	70 mM Tris, pH 8.0; 150 mM sodium chloride; 500 mM imidazole; 6% glycerol; 5 mM β-mercaptoethanol; 0.0174% n-dodecyl-β-D-maltopyranoside
S. aureus SSS buffer A	70 mM Tris, pH 8.0; 150 mM sodium chloride; 20 mM imidazole; 6% glycerol; 5 mM β-mercaptoethanol; 0.174% n-decyl-β-D-maltopyranoside
S. aureus SSS buffer B	70 mM Tris, pH 8.0; 150 mM sodium chloride; 500 mM imidazole; 6 % glycerol; 5 mM β-mercaptoethanol; 0.174 % n-decyl-β-D-maltopyranoside

6.6.8.4 Buffer exchange

Following purification by IMAC, the protein was exchanged into a buffer devoid of any imidazole. Usually, this is the same buffer that will be used for size exclusion chromatography (Section 6.6.8.7). In order to exchange the buffer, proteins were concentrated using an appropriate molecular weight cut off spin concentrator and diluted into the new buffer multiple times.

6.6.8.5 Human rhinovirus 3C protease cleavage

Following buffer exchange, the *Y. pestis* NanT, *P. mirabilis* SSS and *S. aureus* SSS sialic acid transporters required a C-terminal GFP fusion protein and a 6 × Histidine tag to be cleaved off by the human rhinovirus 3C (HRV3C) protease (Hsieh *et al.*, 2010). At a 1 to

12.5 ratio of HRV3C to the protein of interest, cleavage was conducted for 20 hours, at 4 °C, and with gentle agitation. Following cleavage, gravity flow chromatography (Section 6.6.8.6) was conducted prior to size exclusion chromatography (section 6.6.8.7).

6.6.8.6 Gravity flow chromatography

Gravity flow chromatography was used to separate the cleaved protein from both the tag and any protein that failed to cleave. Gravity flow chromatography was carried out using Poly Prep® columns with Ni-NTA agarose resin pre-equilibrated with the same buffer that the protein was exchanged into for cleavage. Cleaved protein was incubated with the pre-equilibrated resin for 1 hour, at 4 °C and with gentle agitation. The protein and resin combination was then loaded into a gravity flow column. The flow through containing the cleaved protein of interest was collected in 1 mL aliquots. The resin was washed with a further three column volumes of buffer to remove any residual cleaved protein.

6.6.8.7 Size exclusion chromatography

Size exclusion chromatography was employed as a final polishing step for all purification procedures described above. Either a HiLoad Superdex 200 16/60 120 mL or Superdex 200 Increase 10/300 GL 24 mL column was employed for size exclusion chromatography. All buffers used for size exclusion chromatography throughout this thesis are listed in Table 2.9. For MRSA *N*-acetylneuraminate lyase, *N*-acetylmannosamine-6-phosphate 2-epimerase, *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase, size exclusion buffer A was selected as the final purification buffer. For the *Y. pestis* NanT, *P. mirabilis* SSS and *S. aureus* SSS sialic acid transporters, the buffers varied depending on the detergent required for each respective protein. Prior to being applied to the column, pooled protein from the previous purification step was concentrated to approximately 1% of the column volume. Eluted fractions corresponding to the 280 nm absorbance peak were analysed by SDS-PAGE for the protein of interest prior to being pooled.

Table 2.9. Size-exclusion chromatography buffers. Components of each buffer are listed.

Buffer	Components
size-exclusion buffer A	20 mM Tris, pH 8.0
Y. pestis NanT size exclusion buffer A	50 mM Tris, pH 8.0; 150 mM sodium chloride; 0.174% n-dodecyl-β-D-maltopyranoside
P. mirabilis SSS size exclusion buffer A	50 mM Tris, pH 8.0; 150 mM sodium chloride; 0.0174% n-dodecyl-β-D-maltopyranoside
S. aureus SSS size exclusion buffer A	50 mM Tris, pH 8.0; 150 mM sodium chloride; 0.174% n-decyl-β-D-maltopyranoside

6.6.9 Protein quantitation

6.6.9.1 The Bradford protein assay to determine protein concentration

The concentration of protein was measured using the Bradford protein assay (Bradford, 1976). Bradford dye reagent was prepared by diluting one part of the concentrate with four parts of MilliQ water. Protein concentrations were obtained from a calibration curve generated in parallel. To construct this calibration curve, bovine serum albumin was serially diluted from 0.5 mg/mL to 0.016 mg/mL in the same buffer that the protein sample was in. Protein samples were diluted in accordance to the midpoint of the calibration curve and prepared in triplicate with 200 µL of the diluted bradford dye reagent and 10 µL of protein. The solution was mixed thoroughly and incubated at room temperature for exactly 5 minutes before measurement of the absorbance at 595 nm. Measurements were taken against a blank consisting of 200 µL bradford dye reagent and 10 µL of buffer.

6.6.9.2 NanoDrop spectrophotometry

Alternatively, protein concentration was measured by UV absorbance at 280 nm using a NanoDrop-1 000 spectrophotometer. The extinction coefficient for each protein being measured was calculated using ExPASy's ProtParam tool (http://web.expasy.org/protparam/). The Beer Lambert law was used to convert the absorbance into concentration.

6.7 X-ray crystallography

6.7.1 Crystallisation

Crystallisation studies were performed using freshly purified preparations of protein, concentrated to 10-20 mg/mL with an appropriate molecular weight cut off spin concentrator. Initial protein crystallisation trials were either performed in-house, or at the Collaborative Crystallisation Centre (C3; http://www.csiro.au/c3/) node of the Commonwealth Scientific and Industrial Research Organisation. All trays were inspected every two to three days for crystal growth.

Initial crystallisation trials for MRSA *N*-acetylneuraminate lyase and *N*-acetylmannosamine-6-phosphate 2-epimerase were performed at C3 using the PACT Suite and the JCSG+ Suite crystal screens (Newman *et al.*, 2005; Newman *et al.*, 2008). These trials were conducted using the sitting drop vapour diffusion method, at 8 °C and 20 °C, with droplets consisting of 150 nL protein solution and 150 nL reservoir solution. In-house optimisation of the initial hits were performed using the hanging drop vapour diffusion method in 24-well Linbro plates, with siliconised cover slips at 8 °C and 20 °C. Hits were first optimised by varying the amounts of protein and reservoir solution; for example, 1.7-2.2 μL of reservoir solution with either 2.0 μL or 2.5 μL of protein solution was equilibrated against 1 mL of reservoir solution. The crystallisation conditions that resulted in diffraction of MRSA *N*-acetylneuraminate lyase and MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase are described in Chapter Two and Chapter Three, respectively.

Crystallisation trials for the *Y. pestis* NanT and *S. aureus* SSS sialic acid transporters were conducted at 20 °C using both vapour diffusion and lipidic cubic phase methods either in-house or at C3. Screens used in-house for vapour diffusion studies included MemGold™, MemGold2™, MemStart™, MemSys™, and The PGA Screen™. The droplets consisted of 200 nL protein solution and 200 nL reservoir solution. All lipidic cubic phase crystallisation was attempted using screens prepared at C3, including the SG1 (Shotgun) Screen, the PACT Suite, Cubic, MemFac, MemGold, and MemGold2. Conditions that the *Y. pestis* NanT sialic acid transporter produced crystals in are described in Chapter Four.

6.7.2 Data collection, processing and structure refinement

6.7.2.1 Data collection

For X-ray data collection, crystals were mounted onto cryo loops of an appropriate size for the crystal of interest, briefly soaked in cryo protectant solution, and flash cooled in liquid nitrogen. Crystals that had been flash cooled were mounted onto the beamline in a cold nitrogen stream at -173 °C. Diffraction data was collected on the MX1 and MX2 beamlines at the Australian Synchrotron.

6.7.2.2 Data processing

Indexing and integration of the data were performed using the program *iMOSFLM* (Battye *et al.*, 2011) or *XDSME* (Kabsch, 2010). The space group was verified by running *POINTLESS* (Evans, 2006) from the CCP4 program suite (Winn *et al.*, 2011). If required, the data was reprocessed in a new space group. Scaling and data reduction were then performed using *SCALA* (Evans, 2006) or *AIMLESS* (Evans, 2006; Evans, 2011), also from CCP4. The resulting intensity data were analysed using *PHENIX XTRIAGE* and the *PHENIX REFLECTION FILE EDITOR* (Adams *et al.*, 2010) was used to generate a new R_{free} set in thin resolution shells. The output allowed the true resolution of the data, the signal to noise ratio, multiplicity, completeness, and bad sections of data to be assessed. An estimate of the number of molecules in the asymmetric unit was obtained using the *MATTHEWS COEFFICIENT* program (Matthews, 1968; Kantardjieff & Rupp, 2003) from the CCP4 suite.

6.7.2.3 Model building and structure refinement

If required, the coordinate file of the search model was prepared using *CHAINSAW* (Stein, 2008) from the CCP4 Suite, which uses a sequence alignment of the search model protein and the target protein to edit the coordinate file. Molecular replacement was used to estimate the phases from a search model using the program *PHASER* (McCoy *et al.*, 2007), also from CCP4. Structure refinement was initially performed using *REFMAC* (Murshudov *et al.*, 2011) in CCP4 and continued thereafter in *PHENIX REFINE* (Adams *et al.*, 2010). The first time that refinement was conducted in *PHENIX REFINE*, simulated annealing

was performed to eliminate model bias. Iterative improvement of both the map and the model was performed using alternate cycles of refinement and residue-by-residue analysis in *COOT* (Emsley & Cowtan, 2004). In *COOT*, manual repositioning of some residues to better fit the electron density was performed. Water molecules were added using the 'add waters' function in *COOT*, and manually assessed for fit to the observed spherical electron density. Ligands were added *via COOT* and modelled into the observed electron density map manually. Coordinate and restraint files for ligands not in the *COOT* library were generated *via SKETCHER* in the CCP4 suite. Following each cycle of refinement, the output was monitored for a decrease in both the residual factor (R_{factor}) and the free R_{factor} (R_{free}). All final images were generated in the program PyMOL.

6.8 Biophysical techniques

6.8.1 Circular dichroism spectroscopy

A Jasco J815 circular dichroism (CD) spectrophotometer was used to record the CD spectra of purified proteins. Wavelength scans between 180 nm and 260 nm were performed on protein samples in an appropriate buffer. Data were collected in a 2 mm path length quartz cuvette, at 20 °C, and in 0.5 nm increments. Buffer data was subtracted from sample data, and data exceeding a voltage of 600 was removed. Data was converted to mean residue ellipticity by multiplying by the conversion factor (Equations 6.3 and 6.4). Resulting data was analysed using the *CDNN* software package (Bohm *et al.*, 1992).

Equation 6.3. $$\textbf{mean residue concentration} = \frac{\textbf{number of amino acids in protein}}{\textbf{molecular weight (g/mol)}} \times \textbf{[sample] (mg/mL)}$$

Equation 6.4 $$\textbf{conversion factor} = \frac{1}{\textbf{mean residue concentration} \times \textbf{path length (mm)}}$$

6.8.2 Differential scanning fluorimetry

Differential scanning fluorimetry was performed with an iQ™5 Multicolour Real Time PCR Detection System (BioRad). Samples consisted of 20 µL 1 × SYPRO® Orange and 80 µL of the relevant protein at 1 mg/mL. Samples were prepared and measured in triplicate in 96 well thick wall plates. Fluorescence was measured for 10 seconds in 0.5 °C increments from 20 °C to 100 °C.

6.8.3 Analytical ultracentrifugation

Sedimentation velocity experiments using analytical ultracentrifugation were performed in an XL-I Analytical Ultracentrifuge (Beckman Coulter). Depending on the protein, the absorbance optical system was employed at wavelengths in the range of 280 to 285 nm. Initial scans were performed at 3 000 rpm to determine the optimal radial settings and wavelength. For these experiments, 400 µL of reference and 380 µL of sample solutions were loaded into 12 mm double sector cells with quartz or sapphire windows, and then

mounted in an An-60 Ti four or eight hole rotor. Proteins were centrifuged at a rotor speed of 50 000 rpm at either 12 °C or 20 °C. Radial absorbance data were collected at a single wavelength without averaging, using a 0.003 cm step size for a total of 80 to 120 scans. Solvent density, solvent viscosity and estimates of the partial specific volume for proteins were computed using the program SEDNTERP (Laue *et al.*, 1992). Data were fitted to a continuous size distribution [c(*s*)] model using the program SEDFIT (Schuck, 2000). This data was then fitted to a continuous mass distribution [c(*M*)] in SEDFIT to calculate the apparent molecular mass of the species observed in the [c(*s*)] distribution.

6.8.4 Small angle X-ray scattering

6.8.4.1 Small angle X-ray scattering measurements

Small angle X-ray scattering (SAXS) data were collected on the SAXS/WAXS beamline, equipped with a Pilatus 1M detector (170 mm × 170 mm, effective pixel size, 172 × 172 μm) at the Australian Synchrotron. The wavelength of the X-rays was 1.0332 Å. A sample detector distance of 1600 mm was used, providing a q range of 0.006-0.4 $Å^{-1}$ (where q is the magnitude of the scattering vector, which is related to the scattering angle (2θ) and the wavelength (λ) as follows: $q = (4\pi/\lambda)\sin\theta$). For all proteins analysed, 50 μL of sample at approximately 10 mg/mL was subjected to size exclusion chromatography on a Superdex 200 5/150 GL column, pre-equilibrated with an appropriate buffer. Data were collected using a 1.5 mm glass capillary, at 20 °C, and at 2 second intervals. Two dimensional intensity plots were radially averaged, normalised to sample transmission, and background subtracted using Scatterbrain software (Australian Synchrotron).

6.8.4.2 Small angle X-ray scattering data analysis

All subsequent SAXS analyses were performed using the ATSAS software package (version 2.3) (Konarev *et al.*, 2006). Guinier plots were analysed using *PRIMUS* (Konarev *et al.*, 2003) to assess data quality. Indirect fourier transform of the data was performed using *GNOM* (Svergun, 1992) to generate the $P(r)$ distribution. Theoretical scattering curves were generated from atomic coordinates and compared with experimental scattering curves using *CRYSOL* (Svergun *et al.*, 1995).

6.9 Enzyme kinetics

6.9.1 Coupled assay for *N*-acetylneuraminate lyase

Kinetic analysis of MRSA *N*-acetylneuraminate lyase was performed using a lactate dehydrogenase coupled assay, as described elsewhere (Devenish & Gerrard, 2009; Timms *et al.*, 2013). The coupled enzyme reaction is illustrated in Figure 6.1. Appropriate controls were performed to ensure that enzyme concentration was proportional to rate, lactate dehydrogenase was in excess, and the temperature selected to run the desired assays were suitable. Controls were performed with 20 mM of the substrate, sialic acid, which was at least 5 × the Michaelis-Menten constant (K_M). The reaction was followed by measuring the consumption of reduced nicotinamide adenine dinucleotide (NADH) by measurement of absorbance at 340 nm as a function of time using the Cary 100 Bio UV/Vis spectrophotometer (Agilent Technologies). Enzyme assays were performed in duplicate or triplicate and initial velocities were reproducible within 10% error.

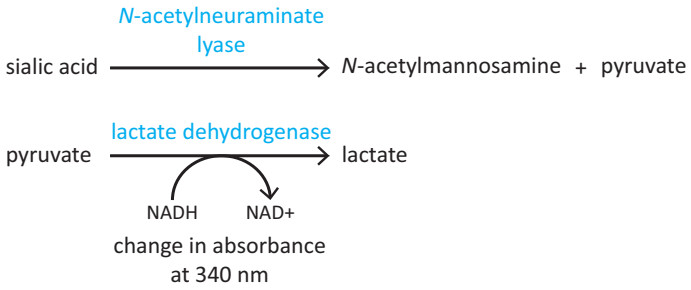


Figure 6.1. The lactate dehydrogenase coupled assay for *N*-acetylneuraminate lyase. As *N*-acetylneuraminate lyase converts sialic acid into *N*-acetylmannosamine and pyruvate, lactate dehydrogenase converts pyruvate into lactate while simultaneously reducing NADH to NAD+.

6.9.1.1 Steady state kinetic analysis of MRSA N-*acetylneuraminate lyase*

Steady state kinetic analysis of MRSA *N*-acetylneuraminate lyase was performed at 30 °C. Reaction mixtures were allowed to equilibrate to 30 °C in the instrument for 10 minutes prior to initiation by addition of 20 µL MRSA *N*-acetylneuraminate lyase. Standard reaction mixtures were made up to 1 mL and contained 100 mM sodium phosphate buffer,

pH 7.6, 0.15 mM NADH, 100 µg LDH, and 0.16 to 40 mM of the substrate [*S*], sialic acid. Initial rate (V_0) data were analysed using OriginPro (version 8.5.1) and fitted to the Michaelis-Menten model (Equation 6.5) to determine the relevant kinetic constants, including the K_M, maximum velocity (V_{max}), and the catalytic turnover number (k_{cat}).

Equation 6.5. $$V_0 = \frac{V_{max}[S]}{K_M + [S]}$$

6.9.1.2 Inhibition studies of MRSA N-*acetylneuraminate lyase*

To evaluate kinetic inhibition of MRSA *N*-acetylneuraminate lyase by the sialic acid alditols, the concentration of the substrate (sialic acid) and a sialic acid alditol mixture were varied against one another. The inhibitor constant (K_i) was determined with 0.16 to 20 mM sialic acid against 0 to 5 mM of the sialic acid alditol mixture. These assays were performed at a physiologically relevant temperature of 37 °C. Initial rate data were analysed using OriginPro and data were fitted to the competitive inhibition model (Equation 6.6), non-competitive inhibition model (Equation 6.7), un-competitive inhibition model (Equation 6.8) and the mixed inhibition model (Equation 6.9). The inhibitor constant for a competitive inhibitor, non-competitive inhibitor, and un-competitive inhibitor is denoted K_{ic}, K_{inc}, and K_{iuc}, respectively. The best fit was determined through consideration of R^2 values and % residuals.

Equation 6.6. $$V_0 = \frac{V_{max}[S]}{K_M\left(1 + \frac{[I]}{K_{ic}}\right) + [S]}$$

Equation 6.7. $$V_0 = \frac{V_{max}[S]}{K_M + [S]\left(1 + \frac{[I]}{K_{inc}}\right)}$$

Equation 6.8. $$V_0 = \frac{V_{max}[S]}{K_M + \left(1 + \frac{[I]}{K_{iuc}}\right)[S]}$$

Equation 6.9. $$V_0 = \frac{V_{max}[S]}{\left(1 + \frac{[I]}{K_{ic}}\right)K_M + \left(1 + \frac{[I]}{K_{iuc}}\right)[S]}$$

6.9.2 Multi-enzyme coupled assay for *N*-acetylmannosamine-6-phosphate 2-epimerase

A new enzyme assay was developed to probe *N*-acetylmannosamine-6-phosphate 2-epimerase function. Kinetic analysis of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase was performed using a multi-enzyme coupled assay; including *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase, the subsequent enzymes of the sialic acid degradation pathway, as well as a phosphoglucoisomerase and glucose-6-phosphate dehydrogenase (Figure 6.2). Briefly, the coupling enzymes work sequentially until glucose-6-phosphate dehydrogenase produces 6-phosphogluconate while simultaneously reducing NADP+ into NADPH. This assay was extended and adapted from the coupled assay described for glucosamine-6-phosphate deaminase (Vincent *et al.*, 2005). Appropriate controls were performed to ensure that enzyme concentration was proportional to rate and all of the coupling enzymes were in excess. The reaction was followed by measurement of absorbance at 340 nm as a function of time using the Cary 100 Bio UV/Vis spectrophotometer. Enzyme assays were performed in duplicate and initial velocities were reproducible within 10% error.

6.9.2.1 Steady state kinetic analysis of MRSA N-*acetylmannosamine-6-phosphate 2-epimerase*

Steady state kinetic analysis of MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase was performed at 25 °C. Reaction mixtures were allowed to equilibrate to 25 °C in the instrument for 10 minutes prior to initiation by addition of 20 µL *N*-acetylmannosamine-6-phosphate 2-epimerase. Standard reaction mixtures were made up to 1 mL and contained 100 mM Tris, pH 8.0, 50 µg *N*-acetylglucosamine-6-phosphate deacetylase, 100 µg glucosamine-6-phosphate deaminase, 50 µg phosphoglucoisomerase and 25 µg glucose-6-phosphate dehydrogenase and an appropriate amount of the substrate, *N*-acetylmannosamine-6-phosphate. The kinetic constants, K_M, V_{max}, and k_{cat} were determined with 0.16 to 5 mM *N*-acetylmannosamine-6-phosphate for the wild type enzyme and 1 to 5 mM *N*-acetylmannosamine-6-phosphate for mutated variants of this enzyme. Initial rate data were analysed using OriginPro and fitted to the Michaelis-Menten model to determine the relevant kinetic constants.

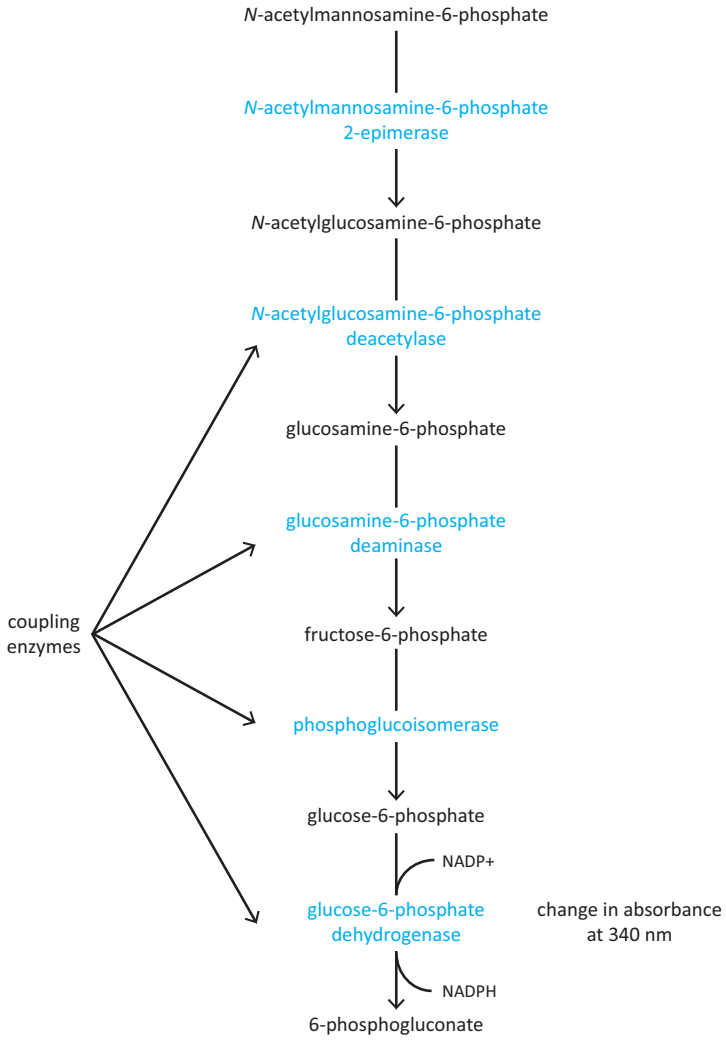


Figure 6.2. The multi-enzyme coupled assay for *N*-acetylmannosamine-6-phosphate 2-epimerase. As *N*-acetylmannosamine-6-phosphate 2-epimerase converts *N*-acetylmannosamine-6-phosphate into *N*-acetylglucosamine-6-phosphate, *N*-acetylglucosamine-6-phosphate deacetylase converts this into glucosamine-6-phosphate. This is then converted into fructose-6-phosphate by glucosamine-6-phosphate deaminase. Next, phosphoglucoisomerase is responsible for the conversion of fructose-6-phosphate to glucose-6-phosphate. Finally, glucose-6-phosphate dehydrogenase converts glucose-6-phosphate 6-phosphogluconate while simultaneously reducing NADP+ into NADPH.

6.9.3 Sodium borohydride reduction

Approximately 1 mg/mL of MRSA *N*-acetylneuraminate lyase and MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase were incubated in the presence of 5 mM of

sialic acid and *N*-acetylmannosamine-6-phosphate respectively, for 30 minutes, at room temperature. The samples were placed on ice and 100 mM sodium borohydride, made up in 100 mM Tris, pH 8.0, was added over 15 minutes. The samples were incubated for a further 30 minutes on ice. The activity of MRSA *N*-acetylneuraminate lyase and MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase was assessed according to their respective coupled assays.

6.10 References

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D, Biological Crystallography, 66*, 213-221.

Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D, Biological Crystallography, 67*, 271-281.

Bohm, G., Muhr, R. & Jaenicke, R. (1992). Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Engineering, 5*, 191-195.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein dye binding. *Analytical Biochemistry, 72*, 248-254.

Devenish, S. R. A. & Gerrard, J. A. (2009). The quaternary structure of *Escherichia coli N*-acetylneuraminate lyase is essential for functional expression. *Biochemical and Biophysical Research Communications, 388*, 107-111.

Drew, D., Lerch, M., Kunji, E., Slotboom, D. J. & de Gier, J. W. (2006). Optimisation of membrane protein overexpression and purification using GFP fusions. *Nature Methods, 3*, 303-313.

Drew, D., von Heijne, G., Nordlund, P. & de Gier, J. W. (2001). Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. *FEBS Letters, 507*, 220-224.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research, 32*, 1792-1797.

Emsley, P. & Cowtan, K. (2004). Coot: model building tools for molecular graphics. *Acta Crystallographica Section D, Biological Crystallography, 60*, 2126-2132.

Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallographica Section D, Biological Crystallography, 62*, 72-82.

Evans, P. R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D, Biological Crystallography, 67*, 282-292.

Hjelm, A., Schlegel, S., Baumgarten, T., Klepsch, M., Wickstrom, D., Drew, D. & de Gier, J. W. (2013). Optimising *E. coli*-based membrane protein production using Lemo21(DE3) and GFP-fusions. *Methods in Molecular Biology, 1033*, 381-400.

Hsieh, J. M., Besserer, G. M., Madej, M. G., Bui, H. Q., Kwon, S. & Abramson, J. (2010). Bridging the gap: a GFP-based strategy for overexpression and purification of membrane proteins with intra and extracellular C-termini. *Protein Science, 19*, 868-880.

Kabsch, W. (2010). Xds. *Acta Crystallographica Section D, Biological Crystallography, 66*, 125-132.

Kantardjieff, K. A. & Rupp, B. (2003). Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Science, 12*, 1865-1871.

Karlin, S. & Altschul, S. F. (1990). Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proceedings of the National Academy of Sciences of the United States of America, 87*, 2264-2268.

Karlin, S. & Altschul, S. F. (1993). Applications and statistics for multiple high-scoring segments in molecular sequences. *Proceedings of the National Academy of Sciences of the United States of America, 90*, 5873-5877.

Konarev, P. V., Petoukhov, M. V., Volkov, V. V. & Svergun, D. I. (2006). ATSAS 2.1, a program package for small-angle scattering data analysis. *Journal of Applied Crystallography, 39*, 277-286.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography, 36*, 1277-1282.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*, 2947-2948.

Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Cambridge, The Royal Society of Chemistry.

Matthews, B. W. (1968). Solvent content of protein crystals. *Journal of Molecular Biology, 33*, 491-497.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography, 40*, 658-674.

Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D, Biological Crystallography, 67*, 355-367.

Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J. L., Stuart, D. I. & Perrakis, A. (2005). Towards rationalisation of crystallisation screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta Crystallographica Section D, Biological Crystallography, 61*, 1426-1431.

Newman, J., Pham, T. M. & Peat, T. S. (2008). Phoenito experiments: combining the strengths of commercial crystallisation automation. *Acta Crystallographica Section F, Structural Biology and Crystallographic Communications, 64*, 991-996.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical Journal, 78*, 1606-1619.

Stein, N. (2008). CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *Journal of Applied Crystallography, 41*, 641-643.

Suzuki, H., Tabata, K., Morita, E., Kawasaki, M., Kato, R., Dobson, R. C., Yoshimori, T. & Wakatsuki, S. (2014). Structural basis of the autophagy-related LC3/Atg13 LIR complex: recognition and interaction mechanism. *Structure, 22*, 47-58.

Svergun, D., Barberato, C. & Koch, M. H. J. (1995). CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography, 28*, 768-773.

Svergun, D. I. (1992). Determination of the regularisation parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography, 25*, 495-503.

Timms, N., Windle, C. L., Polyakova, A., Ault, J. R., Trinh, C. H., Pearson, A. R., Nelson, A. & Berry, A. (2013). Structural insights into the recovery of aldolase activity in *N*-acetylneuraminic acid lyase by replacement of the catalytically active lysine with

gamma-thialysine by using a chemical mutagenesis strategy. *Chembiochem: a European Journal of Chemical Biology, 14*, 474-481.

Vincent, F., Davies, G. J. & Brannigan, J. A. (2005). Structure and kinetics of a monomeric glucosamine 6-phosphate deaminase: missing link of the NagB superfamily? *Journal of Biological Chemistry, 280*, 19649-19655.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D, Biological Crystallography, 67*, 235-242.

Appendix

Methicillin resistant *Staphylococcus aureus* (MRSA) *N*-acetylneuraminate lyase

ATGAACAAAGATTTAAAAGGTTTATATGCAGCGTTACTCGTTCCTTTTGATGAAAATGGT
CAAGTAAATGAACAAGGACTTAAACAAATTGCTCAAAATGCCATTGAAACTGAAGAATTA
GACGGTCTTTATGTAAATGGTAGCTCAGGCGAAAACTTTTTATTAAATACAGAACAGAAG
AAGCAAGTTTTCAAAGTTGCCAAAGAGGCAGTTGGGGATAAAGTGAAATTGATTGCTCAA
GTAGGTTCGTTAGATTTAAATGAAGCCATTGAACTCGGAAAATATGCCACAGAACTCGGT
TATGATGCACTTTCTGCCGTAACACCATTCTACTATCCATTTACTTTTGAAGAAATTAGA
GATTACTATTTCGATATTATCGAAGCAACACAGAACAATATGATTATTTATGCAATACCA
GATTTAACAGGTGTGAATATTTCTATTGAACAATTCAGCGAGCTTTTTAATCACGAAAAA
ATTGTTGGTGTTAAATATACAGCGCCAAACTTCTTCCTACTCGAACGTATTAGAAAAGCA
TTCCCAGACAAATTAATCTTATCTGGCTTTGATGAAATGTTAGTTCAAGCTACGATTTCT
GGCGTAGATGGTGCAATTGGTTCTACATATAACGTTAATGGTCGCCGCGCTAGAAAAATC
TTCGACTTAGCAAGACAAGGTCAAATTCAAGAAGCTTACCAACTACAACATGATTCAAAC
GACATCATCGAAACAGTGTTATCAATGGGGATTTATCCAACATTGAAAGAAATACTACGT
CACCGTGGCATTGATGCAGGATTACCGAAACGTCCATTCAAACCTTTTAACGAAGCACAT
CGACAAACATTAGATCAACTCATTGCAAAATACGATTTATAG

MRSA *N*-acetylmannosamine-6-phosphate 2-epimerase

ATGTTACCACATGGATTAATAGTATCTTGTCAGGCACTACCAGATGAACCATTGCATTCA
TCTTTTATTATGTCGAAAATGGCATTAGCTGCGTATGAAGGTGGTGCTGTTGGTATTCGC
GCAAATACTAAGGAAGACATTTTAGCAATTAAAGAAACGGTAGATTTACCAGTTATTGGC
ATTGTGAAACGTGACTATGATCACTCAGATGTTTTCATTACTGCAACGTCAAAAGAAGTT
GATGAACTGATAGAAAGCCAATGTGAAGTCATTGCATTGGATGCAACGTTACAGCAACGT
CCGAAAGAAACGTTAGACGAATTAGTATCATATATTAGAACACATGCACCGAACGTTGAA
ATCATGGCTGATATCGCGACCGTTGAAGAAGCTAAAAATGCCGCACGACTTGGCTTTGAT
TATATTGGCACGACGTTACATGGCTATACTAGTTATACGCAAGGACAATTACTTTATCAA
AATGACTTCCAATTTTTAAAAGATGTACTACAAAGTGTTGATGCAAAAGTTATTGCGGAA
GGTAATGTCATTACACCGGATATGTATAAACGTGTGATGGACTTAGGCGTTCATTGTTCA
GTCGTTGGTGGTGCGATAACACGACCAAAAGAAATTACGAAACGTTTTGTTCAAATTATG
GAAGATTAA

NanE forward primer

5′-CATATGTTACCACATGGATTAATAGTATCTTGTCAGGC-3′

NanE reverse primer

5′-AAGCTTTTAATCTTCCATAATTTGAACAAAACGTTTCG-3′

MRSA *N*-acetylglucosamine-6-phosphate deacetylase

ATGTCAGAATTAATTATATATAACGGCAAAGTTTATACTGAAGATGGCAAAATCGATAAT
GGTTACATTCATGTGAAAGATGGACAGATTGTTGCAATTGGAGAAGTGGATGATAAAGCA
GCAATTGATAATGATACGACAAATAAAATTCAAGTGATTGATGCTAAAGGTCATCATGTA
TTACCAGGTTTTATTGATATACATATTCATGGTGGTTATGGTCAAGATGCAATGGATGGG
TCATACGATGGCTTAAAATATCTATCCGAAAATTTGTTGTCTGAAGGGACGACATCATAC
TTGGCCACTACAATGACGCAATCGACTGATAAAATAGATAATGCACTTACAAATATTGCT
AAATATGAAGCGGAGCAAGATGTTCACAATGCAGCGGAAATTGTAGGTATACATTTAGAA
GGACCATTTATATCTGAAAATAAAGTTGGTGCTCAACATCCGCAATACGTTGTACGCCCA
TTTATCGATAAAATTAAACATTTTCAAGAGACTGCTAACGGATTAATAAAGATTATGACG
TTTGCACCTGAAGTTGAAGGTGCAAAAGAAGCGCTTGAAACGTATAAAGATGACATTATT
TTTTCAATTGGTCATACAGTAGCAACATACGAAGAAGCAGTTGAAGCTGTTGAGCGAGGA
GCTAAACATGTCACGCATTTATATAATGCAGCGACGCCATTCCAACATAGAGAACCAGGT
GTTTTTGGAGCAGCATGGTTGAATGATGCTCTACATACCGAAATGATTGTTGATGGCACT
CATTCTCATCCGGCATCGGTTGCAATTGCTTACCGTATGAAAGGTAATGAACGTTTTTAT
TTAATTACCGATGCAATGCGTGCAAAAGGTATGCCTGAAGGAGAATATGATTTGGGTGGA
CAAAAAGTAACTGTTCAATCGCAACAAGCACGTCTTGCAAATGGTGCGCTTGCTGGTAGT
ATTTTAAAAATGAATCATGGGTTACGTAACTTAATATCATTTACAGGTGATACATTAGAT
CATTTATGGCGAGTAACAAGTTTAAATCAAGCCATTGCATTAGGTATCGATGATAGAAAA
GGTAGTATTAAAGTAAATAAGGATGCAGATCTTGTTATTCTAGATGATGATATGAATGTA
AAATCTACAATAAAACAAGGCAAGGTTCACACATTTAGCTAA

MRSA glucosamine-6-phosphate deaminase

ATGAAAGTATTAAACTTAGGATCGAAAAAACAAGCATCATTCTATGTTGCATGTGAGTTA
TATAAAGAGATGGCATTTAATCAGCACTGTAAACTAGGTTTAGCAACTGGTGGTACAATG
ACAGATTTGTATGAACAACTTGTTAAGTTATTAAATAAAAATCAGTTAAACGTAGACAAT
GTATCCACGTTTAATTTAGACGAATATGTAGGTTTAACCGCATCACATCCGCAAAGTTAT
CACTATTATATGGATGACATGCTTTTCAAACAATATCCTTATTTTAATAGAAAGAACATT
CATATTCCAAATGGAGATGCCGATGATATGAATGCGGAAGCGTCAAAATATAATGACGTT
TTAGAACAACAAGGTCAACGTGATATTCAAATTTTAGGTATTGGTGAAAATGGTCATATT
GGATTTAATGAACCTGGTACGCCGTTTGATAGCGTTACTCATATCGTTGATTTGACTGAA
AGTACTATTAAGGCTAATAGTCGATATTTTAAAAACGAAGATGATGTTCCAAAGCAAGCC
ATTTCGATGGGACTTGCTAATATTCTTCAAGCCAAACGTATCATTTTACTCGCATTTGGT
GAAAAGAAACGTGCTGCTATTACACATTTATTAAATCAGGAAATTTCTGTTGATGTTCCA
GCCACATTACTTCACAAACACCCGAATGTTGAGATATATTTAGACGACGAAGCCTGCCCG
AAAAATGTTGCGAAAATTCATGTCGATGAAATGGATTGA

***Yersinia pestis* sugar proton symporter (NanT) for sialic acid transport**

ATGAGTATTTCAGTAGGTCCATCTCGTGAAGATAAACCATTATCGGGTGGCGCTAAACCA
CCCCGTTGGTACAAACAACTTACCCCGGCGCAATGGAAGGCCTTTGTTGCCGCTTGGATC
GGTTATGCCCTGGATGGCTTTGACTTTGTTCTGATTACTCTGGTTCTGACCGATATTAAA
CAAGAATTTGGCCTGACACTGATTCAGGCGACCAGCCTGATTTCTGCTGCCTTCATCTCA
CGCTGGTTTGGTGGGTTGGTACTGGGCGCGATGGGGGATCGCTATGGCCGTAAACTGGCC
ATGATCACCAGTATTGTGTTGTTCTCCTTCGGTACGTTGGCCTGTGGCTTAGCACCTGGC
TACACCACGCTGTTTATTGCTCGCTTGATTATCGGTATTGGCATGGCGGGTGAGTATGGT
TCCAGCTCGACCTATGTGATGGAAAGCTGGCCTAAAAACATGCGTAATAAAGCCAGTGGC
TTCCTGATTTCTGGCTTCTCTATCGGTGCGGTACTCGCGGCGCAAGCCTACAGCTACGTG
GTGCTCGCATTTGGTTGGCGTATGTTGTTCTACATTGGATTATTGCCAATTATCTTTGCA

CTGTGGTTGCGTAAAAACCTACCGGAAGCAGAGGACTGGGAAAAGGCACAAAGTAAGCAG
AAAAAAGGTAAACAGGTCACTGACCGGAATATGGTGGATATTCTGTATCGCAGTCACCTC
AGTTATCTGAATATTGGCCTGACGATATTTGCCGCTGTCTCACTTTACCTCTGTTTTACT
GGCATGGTCTCGACGTTGCTGGTAGTGGTTCTCGGTATTCTTTGCGCTGCAATATTTATC
TATTTTATGGTTCAAACCAGTGGCGATCGCTGGCCTACGGGCGTCATGCTGATGGTCGTG
GTGTTCTGTGCGTTCCTCTATTCTTGGCCGATCCAGGCGTTGTTACCGACCTACCTGAAA
ATGGATCTCGGCTATGACCCACACACCGTAGGCAATATATTGTTCTTCAGTGGTTTTGGT
GCGGCCGTGGGTTGTTGTGTTGGCGGTTTCCTTGGCGATTGGTTGGGTACCCGCAAAGCC
TATGTGACCAGTTTGCTGATATCACAGCTCTTGATCATCCCGCTGTTTGCCATCCAAGGC
AGCAGTATTTTGTTCCTGGGGGGATTACTGTTCTTACAACAGATGCTGGGGCAGGGGATT
GCGGGCCTGTTGCCGAAACTGCTGGGCGGTTATTTTGATACCGAACAGCGAGCCGCAGGA
CTGGGCTTTACCTACAACGTCGGCGCATTGGGAGGGGCATTGGCACCCATACTGGGGGCA
TCGATTGCTCAACATCTCAGTTTAGGCACCGCGTTGGGATCGCTCTCTTTCAGTCTGACA
TTCGTGGTGATCCTACTGATTGGTTTTGATATGCCATCCCGTGTACAGCGTTGGGTCCGC
CCATCAGGTTTACGGATGGTGGATGCCATCGATGGCAAACCATTCAGTGGTGCCATTACG
GCCCAGCATGCGCGAGTAGTGACACAGAAATAA

***Proteus mirabilis* sodium solute symporter (SSS) for sialic acid transport**

ATGCAATTACATGATTTTGGTTTTATTAACTATGCAGTGCTGTTTGGCTATTTGGCAGCT
ATGTTACTCGTGGGTGTTTATTTTTCTAAACGTCAAAAAACTGCGGATGACTATTTCCGT
GGTGGAGGGCGAGTTCCTGGGTGGGCCGCTGGTGTGTCTGTATTTGCGACCACATTAAGC
TCGATTACATTTATGTCCATTCCTGCTAAAGCTTATACTTCTGATTGGACCTTTATTATT
GGACAGTATTTAGCTATTGCAATTTTACCTTTAGTTTTTTATTTCTATATTCCTTTTTTT
AGAAAATTAAAAATCACCTCCGCTTATGAATATCTAGAAGCTCGTTTTGATGTTCGTAGT
CGTTTATTTGCTAGCCTTTCATTTATGTTATTCCATATTGGGCGTGTTGCTATTATTACT
TATTTAACCGTACTAGCTTTACGTCCTTTTATGGGGATTGATCCGGTCGTGTTAATCGTA
TTAATAAGCTTATTATGCATTATTTATACTTGGATGGGGGGCATTGAAGGCGTTATTTGG
ACGGATGTTATTCAAGGTTTATTACTTTCTGGTGGTGCTGTACTGATCTTTATTATGATC
TGTTTTAAAGTTGATGGTGGTATTAGCGAAATTTTTACTACTACAGCACAAGCAGATAAA
TTCTTTCCTACCACACAATGGCGCTGGAGTTGGACAGATAGCACTATTCCTGTATTAATG
ATTGGGTTTTTATTTGCCAATATTCAGCAATTTACTGCTAGTCAGGATGTTGTGCAGCGC
TATATCGTTACCGATTCTATTAAAGAAACTAAACGTACATTGATAACTAACGCTAAATTA
GTGGCTATTATTCCTATTTTCTTTTTTGCCATCGGCTCTGCATTGTTTGTTTACTACCAA
CAAAACCCAAGTTTATTACCCGCAGGATTTAATACGGGGGGAATTTTACCACTATTTATC
GTGACGGAAATGCCTATCGGTATTGCAGGGTTAATTATTGCAGCTATTTTTGCTGCTGCA
CAATCAAGTATTTCTAGTAGTTTAAATAGTATTTCAAGTTGCTTTAATTCTGATATTTAT
ACTCGTCTTAGTAAATCATCCCCTAGTCCTGAACAAAAAATGAAAGTCGCAAAGTTAGTT
ATTATTGTGGCAGGGATATTCAGTAGTTTAGCGGCAATTTGGTTAGTTTTATCCGATGAA
GCTGAAATTTGGGATGCATTTAATAGCCTTATAGGTCTTATGGGCGGACCAATGACAGGC
CTATTTATGCTGGGAATTTTTGTAAAACGTGCTAATGCCGGTAGCGCCGTTGTTGGGATC
ATCGTGAGTATTATTGCAGTATTGGCTGCACGTTATGGCAGTGATCTTAACTTCTTCTTC
TATGGAGTCATTGGTTCAATGTCAGTGGTGATTGCAGGAACCATTACGGCACCACTTTTC
GCACCAGCAAAACAGCTTTCCTTAGATGACAGTGAAACATCAGAGAACTAA

***S. aureus* SSS for sialic acid transport**

ATGAAAGAAGTTGGTTTTGGCACCCTGAATTGGGTTGCAGTTATTATCTATCTGCTGGCC
ATGCTGTTTATCGGTGTGTATTTTACCAAACGTGCAAGCCAGAGCACCAATAGCTTTTTT

ACCGCAAGCGGTCGTCTGCCGAGCTGGGTTGTTGGTTTTAGCATTTATGCAACCACCCTG
AGCGCAATTACCTTTATGAGCACACCGGAAAAAGCATTTCTGACCGATTGGAGCTATATT
GCAGGTAATATTGCCATTGTTGCCATTATTCCGCTGCTGATCTATTTCTATGTGCCGTTC
TTCAAAAAACTGAAAGTTACCAGCGCCTATGAATATCTGGAAGCACGTTTTGGTCCGAGC
ATTCGTGTTATTGGTAGCCTGCTGTTTGTTGTTTATCATCTGGGTCGTGTTGCCATCGTT
ATTTATCTGCCGACCCTGGCAATTACCAGCGTTAGCGATATGAATCCGTATATTGTTGCA
AGCCTGGTTGGTCTGCTGTGTATTCTGTATACCTTTCTGGGTGGTTTTGAAGGTGTTGTT
TGGAGCGATTTTATTCAGGGCGTTATTCTGCTGGGTGGTGCACTGGTTATCATTATTCTG
GGTGTGATGAACATCAAAGGTGGCTTTGGCACCGTTTTTGCAGATGCAATTGAACATAAA
AAACTGATCAGCGCAGACAACTGGAAACTGAATACCGCAGCAGCAGCAATTCCGATTATC
TTTCTGGGCAACATTTTCAACAACCTGTATCAGTATACCGCCAGCCAGGATGTTGTTCAG
CGTTATCAGGCAAGCGATAGCCTGAAAGAAACCAATAAAAGCCTGTGGACCAATGGTATT
CTGGCACTGATTAGCGCACCGCTGTTTTATGGTATGGGCACCATGCTGTATAGCTTTTAT
GCACATGAAGCAGTTCTGCCGAAAGGTTTTAATACCAGCAGCGTTGTTCCGTATTTTATC
CTGACCGAAATGCCTCCGTTTGTTGCAGGTCTGCTGATTGCAGCAATTTTTGCAGCAGCA
CAGAGTACCATTAGCAGCAGCCTGAATAGCATTAGCGCATGTATTAGCATCGATATCAAA
CAGCGCTTTTTTGGTAAAGGTAGCGAACGTCATGAAGTGAATTTTGCCCGTTTCATTATT
ATCATTGCCGGTATCTTTGGCTTTGGTATGAGCCTGTATCTGATTGCCAGTAATAGCAAT
GATCTGTGGGACCTGTTTCTGTTTGTTACCGGTCTGTTTGGTGTTCCGCTGGCAGGCGTT
TTTGCCGTTGGTATTTTTACAAAACGTACCAATACCTTTGGCGTGATTTGTGGTCTGATT
CTGGGTATTATCTTCGCCTATGTTTATAATGGTGTGGGCAAAGGTAATAGCCCGTTTTAT
GTTAGCACCATCAGCTTTACCGTTGCCTTTGTTTTTGCATATATCCTGAGCTTTATTGTG
CCGAGCAAACACAAAAAAGATATTACGGGTCTGACCATCTTTGAAAAAGATAAACCGAGC
ACCTACATTAGCAAAACCGCAACCAAAAAATGA